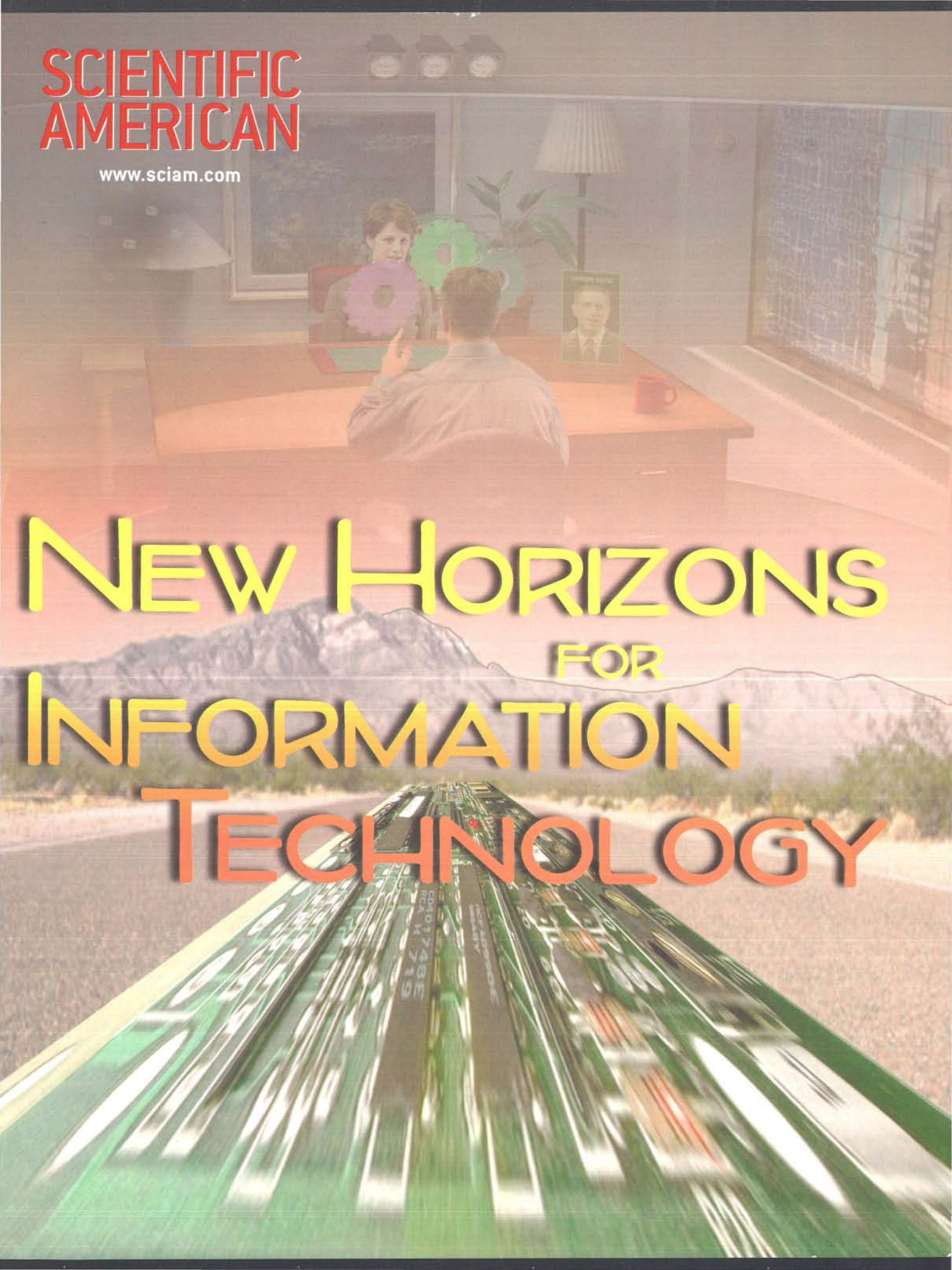


SCIENTIFIC
AMERICAN

www.sciam.com



NEW HORIZONS FOR INFORMATION TECHNOLOGY

FROM THE EDITORS

Dear SCIENTIFIC AMERICAN Subscriber:

Moore's Law is cited in articles about computing so often and so emphatically that a naïve reader might mistakenly think it was an inviolable principle of nature, like Newton's Law of Gravitation or the Laws of Thermodynamics. It is actually something more remarkable. In the mid-1960s Gordon Moore, the founder of Intel Corp., observed that roughly every two years manufacturers doubled the number of transistors and other components packed onto microprocessors, and he suggested that this trend would continue. From that accurate prediction of geometric increase flowed all the amazing advances in speed, complexity and capability exhibited by modern computers and other electronics.

But skeptics have always questioned whether that accelerating pace could be maintained indefinitely: surely, at some point, limitations of physics or economics would throw on the brakes. So far, however, despite the obstacles, engineers have kept finding ways to make transistors smaller. Moore's Law is thus not so much a fact of nature as it is a vote of confidence in human ingenuity.

That same ingenuity is on display in this collection of articles from Scientific American. All of them are about pushing back the horizon of what seems achievable through computing. Some of these articles, as you might expect, are indeed about the challenges of making semiconductor circuitry more and more compact. Most of these articles, however, are more concerned with inventive ways of applying information technology.

The spirit of invention is the key to everything from novel chip designs and networking schemes to ways of turning coalitions of simple PCs into functional supercomputers. Some engineers are finding electronic replacements for paper; others are creating virtual-reality replacements for travel. While some hackers plot to bring down the Web, the Web's creator is envisioning a better, more capable successor to it. And a few researchers are even investigating the surprisingly difficult problem of designing machines capable of making copies of themselves.

The cleverness on display in these Scientific American articles more than guarantees that even if Moore's Law someday falters, the information age will continue to deliver miracles.

Sincerely,



John Rennie, Editor in Chief
SCIENTIFIC AMERICAN

NEW HORIZONS FOR INFORMATION TECHNOLOGY

CONT

February 2002

4



THE NETWORK IN EVERY ROOM BY W. WAYT GIBBS

THANKS TO INGENUOUS
ENGINEERING, COMPUTERS
AND APPLIANCES CAN NOW
COMMUNICATE THROUGH
THE ELECTRICAL POWER
WIRING IN A HOUSE.

January 2002

10



A VERTICAL LEAP FOR MICROCHIPS BY THOMAS H. LEE

ENGINEERS HAVE DISCOVERED
HOW TO PACK MORE
COMPUTING POWER INTO
MICROCIRCUITS: BUILD THEM
VERTICALLY AS WELL AS
HORIZONTALLY.

November 2001

18



THE ELECTRONIC PAPER CHASE BY STEVE DITLEA

COMPANIES WITH COMPETING
TECHNOLOGIES RACE TO
CREATE DIGITAL "PAPER" THAT
COMBINES THE BEST FEATURES
OF PRINT AND COMPUTER
DISPLAYS.

October 2001

24



CODE RED FOR THE WEB BY CAROLYN MEINEL

THE RECENT INTERNET WORM
ASSAULT MAY PRESAGE MORE
VIRULENT COMPUTER
SABOTAGE IN THE FUTURE.

September 2001

33



THE INCREDIBLE SHRINKING CIRCUIT BY CHARLES M. LIEBER

RESEARCHERS HAVE BUILT
NANORESISTORS AND
NANOWIRES. NOW THEY HAVE
TO FIND A WAY TO PUT THEM
TOGETHER.

August 2001

40



GO FORTH AND REPLICATE

BY MOSHE SIPPER AND
JAMES A. REGGIA

BIRDS DO IT, BEES DO IT, BUT
COULD MACHINES DO IT? NEW
SIMULATIONS SUGGEST YES.

August 2001

50



THE DO-IT-YOURSELF SUPERCOMPUTER

BY WILLIAM W.
HARGROVE, FORREST M.
HOFFMAN AND THOMAS
STERLING

NETWORKS OF ORDINARY
DESKTOP PCs CAN TACKLE
FIENDISHLY HARD JOBS ON THE
CHEAP.

July 2001

58



HOW TO BUILD A HYPERCOMPUTER

BY THOMAS STERLING

HYBRID DESIGNS ACCELERATE
COMPUTING TO MORE THAN A
QUADRILLION OPERATIONS PER
SECOND.

Copyright © 2001, 2002, 2003 by Scientific American, Inc. All rights reserved. Printed in the U.S.A. No part of this reprint may be reproduced by any mechanical, photographic or electronic process, or in the form of a recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without prior written permission of the publisher. The trademark and trade name "SCIENTIFIC AMERICAN" and the distinctive logotype pertaining thereto are the sole property of and are registered under the name of Scientific American, Inc. Page numbers and internal references may vary from those in the original issues

April 2001

66



VIRTUALLY THERE

BY JARON LANIER

TELE-IMMERSION, A NEW
COMMUNICATIONS MEDIUM,
ALLOWS PEOPLE WHO ARE
APART TO FEEL AS THOUGH
THEY ARE IN THE SAME ROOM.

THE NETWORK IN EVERY ROOM

BY W. WAYT GIBBS

Information is power, or so it is often said. Now the reverse is also true: a new technology coming to market this winter allows information to be communicated at high speed over the existing power lines in a building. Connecting all the network-aware devices in a house—at first just computers and printers but in the future telephones, entertainment components and even appliances—will soon be a simple matter of plugging them in. If one of the devices on the home network is a cable modem or DSL (digital subscriber line) router, then all the outlets in the house can tap into the Internet as well as the power grid.

Engineers have suspected for years that there might be clever ways to force into dual duty all the uncountable miles of copper wiring that reach into nearly every room in the industrial world. Indeed, the first U.S. patent on a method to use the power wires for communications was issued in 1899. But many early attempts slammed into immovable technical obstacles. After several embarrassing failures by well-known firms such as Nortel Networks and Siemens AG, a few European utilities have at last begun to offer their customers phone and Internet service over power lines. The American, Canadian and Japanese power grids use a different design that makes it far too costly for utilities to compete with DSL and cable, however [see box on page 6].

Meanwhile other technologies have arrived that can link up machines inside a building without stringing cables through the walls. The new power-line networking standard will have to

compete with HomePNA, for example, which works through telephone jacks. Another challenger, known formally as 802.11b and colloquially as Wi-Fi, communicates via radio waves.

Both HomePNA and Wi-Fi have enjoyed only a lukewarm reception among homeowners, says Kurt Scherf, who tracks the home networking industry for Parks Associates in Dallas. Of the 26 million or so households in the U.S. that have multiple computers, Scherf says, "there are probably 5.5 million already networked, and about 80 to 90 percent of those use Ethernet." At 100 megabits a second, an Ethernet network transmits at 10 times the speed of HomePNA and Wi-Fi, but it must use special cables.

Wi-Fi is still relatively expensive (about \$400 for a base station and two adapters), and many potential users are worried about neighbors eavesdropping on the e-mails they send and the Web sites they browse. Rightly so, Scherf says: "There is an encryption option for almost all wireless network devices, but most users don't know how to turn it on." The biggest limitation of HomePNA is that apartments and older homes often have only one or two phone jacks.

Power outlets, in contrast, are ubiquitous. And practically any machine you'd want to connect needs electricity anyway. Because power-line communications equipment doesn't require radio transducers, it can be less expensive as well as more secure than Wi-Fi. The trick is making it reliable and fast. Previous standards for sending data over electrical wiring—X-10,

Thanks to ingenious engineering, computers and appliances can now communicate through the electrical wiring in a house



Communicating over power lines is like transcribing a symphony played over stadium loudspeakers, during the Super Bowl, while wearing earmuffs.

CEBus and LonWorks have been the most popular—top out at roughly 10 kilobits a second. HomePlug Powerline Alliance, a consortium of some 90 electronics and computer companies, knew it had to find a way to increase that speed 1,000-fold.

A Silk Purse from a Spider's Web

THAT CHALLENGE APPEALED to Larry W. Yonge, vice president for research at Intellon in Ocala, Fla. A communications engineer, Yonge had founded a satellite dish company that boomed for several years but ultimately succumbed to competition from cable. When he joined Intellon in 1996, his sole task was to solve the power-line network problem. "It wasn't immediately obvious that it could be done," he recalls. "We had to design a whole system from scratch."

Even engineers used to dealing with interference say that power lines are a truly horrendous medium for communica-

tions. "We have designed a very efficient system to transmit electrical energy at 60 hertz," says William E. Blair of the Electrical Power Research Institute in Palo Alto, Calif. "But if you take the same wires and try to put communications signals on them, it's a disaster." Rather like trying to transcribe the exact score of a symphony—as it is played over the loudspeakers at a packed football stadium, on a windy night, during the final minutes of a tied Super Bowl game, while wearing earmuffs.

For the noise of the crowd, substitute hair dryers, light dimmers, fluorescent lamps, vacuum cleaners and a cacophony of similar appliances that send spikes of static coursing back through the wires. In much the same way that a public-address system distorts the sound of music, the house wiring itself filters out some frequencies preferentially, altering the signal in erratic ways. The data stream ricochets off unused outlets and other sudden changes in the resistance of the wire, setting up standing waves and resonating feedback.

Like earmuffs, the home's main circuit breaker dampens signals because it splits the wiring into nearly separate sections. In newer houses that have so-called dual-phase power, a message put on the "hot" wire often has to travel up the street to a distribution transformer hanging on a utility pole, then zip back on the "ground" wire in order to reach a different part of the building. By the time it reaches the target outlet, the signal can be faint indeed.

Compounding these problems, the shape and size of the power network change unpredictably whenever a switch is thrown. Turn on the lights or plug in an extension cord, and the voltage, resistance and current fluctuate.

To top it all off, Blair adds, the long, looping electrical circuits act as antennas: they pick up a variety of radio signals that can interfere with the data signal. But if the data signal is boosted too much to overcome the noise, the wires will also transmit radio signals of their own. "So you get interference in everything else," Blair explains, "like a buzz in your stereo and static on your television."

Yonge and his assistants worked around these problems in various ways. They decided to use much higher frequencies than anyone had tried before, above four megahertz. Appliances generate less noise in that part of the electromagnetic spectrum. And multimegahertz signals tend to pass relatively smoothly from one side of the circuit breaker (or distribution transformer) to the other. Off go the earmuffs.

The engineers also elected to use a very wide swath of spectrum, from 4.5 megahertz up to 21 megahertz, which they then divided into 84 channels. Much like the technique used by digital cellular phones, this approach allows devices to send data with less power and thus less pollution of the radio environ-

Power, Gas and Web

Utilities are becoming ISPs, but not in North America

Beyond its use in home networks, electrical wiring can now connect entire neighborhoods to the Internet—but only in certain parts of the world. That is because it is not just the shape of a power plug that differs from country to country; the structure of the power grid differs as well. In most of Europe, Asia and South America, a single large transformer serves hundreds of buildings. The U.S., Canada and Japan, among others, use smaller pole-top transformers that connect to only three to six homes.

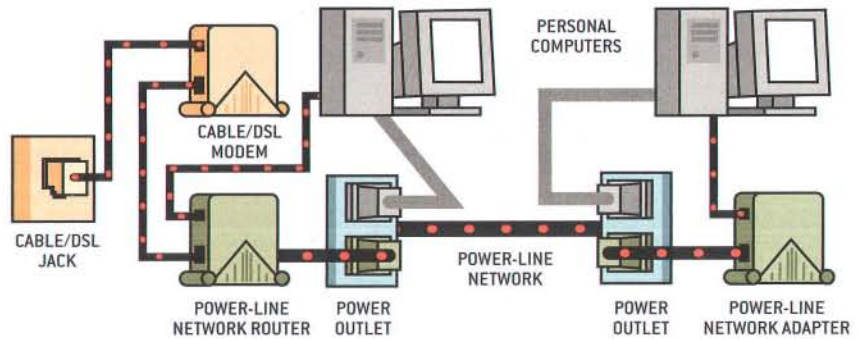
Unfortunately for utilities that want to provide Internet services via their wires, transformers act like big filters that erase information-bearing ripples from the current. "You can put a bypass amplifier around each transformer," says William E. Blair of the Electric Power Research Institute in Palo Alto, Calif., "but it breaks the economics to put dozens on each circuit, as you would have to do in the U.S. Europeans can have one very good bypass amplifier at a substation that routes the signals around the transformer to 200 or 300 homes. That's cost-effective."

Indeed, last summer RWE Powerline, a utility subsidiary based in Essen, Germany, began offering two-megabit-a-second service for about 35 euro (\$31) a month. It uses adapters made by the Swiss firm Ascom, which reports that Freiburg Electric Enterprises in Switzerland is also rolling out Internet service and that field trials are under way in nine other European countries, Hong Kong, Singapore and Brazil.

—W.W.G.

PLUGGING INTO A HOME NETWORK: JUST ADD POWER

THE FIRST POWER-LINE networking products to hit the market early in 2002 will connect only to computers, printers, scanners and other devices that accept USB or Ethernet connections. But the HomePlug standard will work as well with future household appliances modified to communicate through their power cords.



1 By connecting a power-line router to her DSL modem, Sue extends high-speed Internet access to every wall.

2 Bobby can now use the PC in his room to surf the Web for school research and sports news.

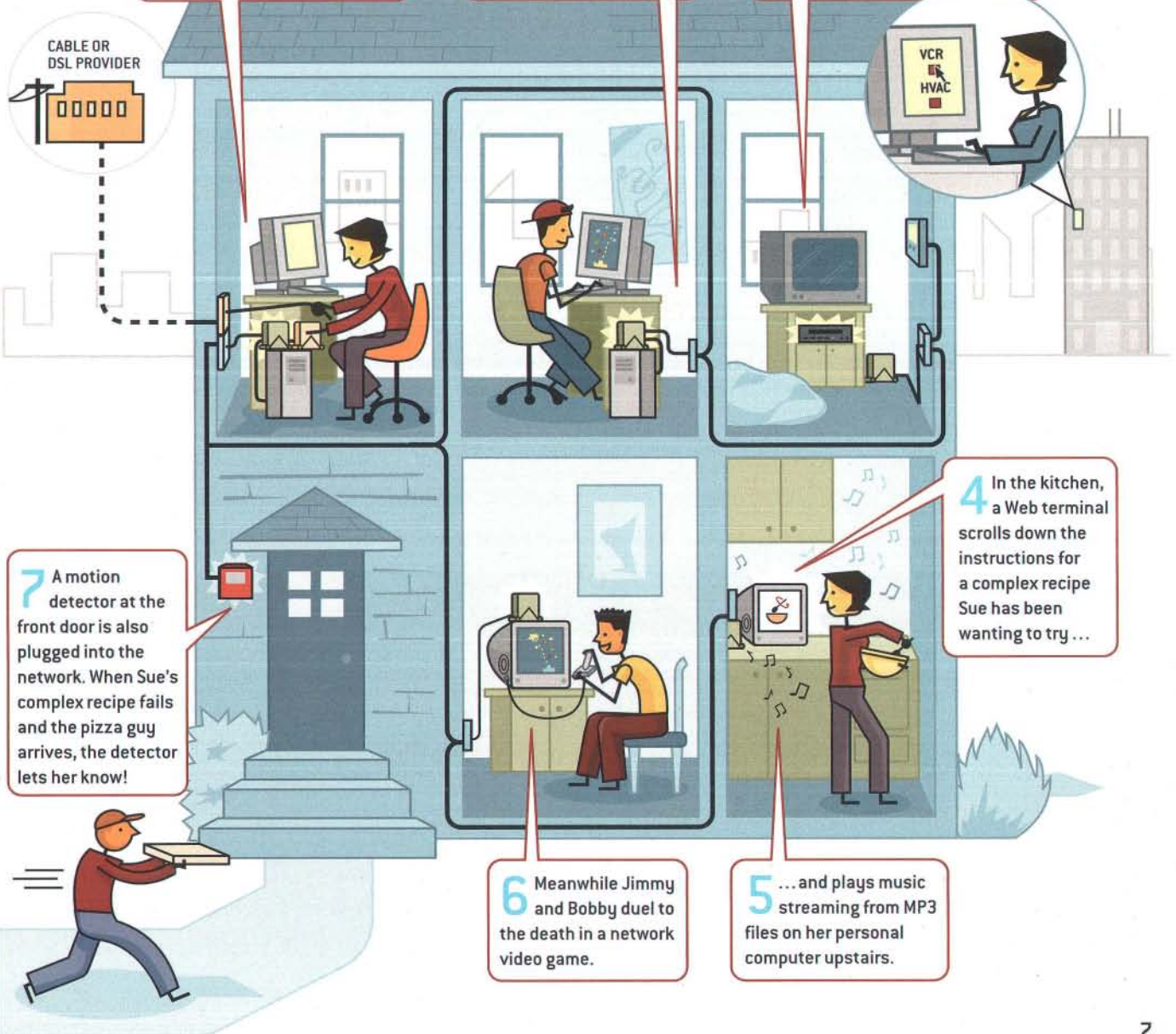
3 From her office, Sue enters a password and clicks on a Web page to turn on the VCR and lower the thermostat in her house.

4 In the kitchen, a Web terminal scrolls down the instructions for a complex recipe Sue has been wanting to try ...

7 A motion detector at the front door is also plugged into the network. When Sue's complex recipe fails and the pizza guy arrives, the detector lets her know!

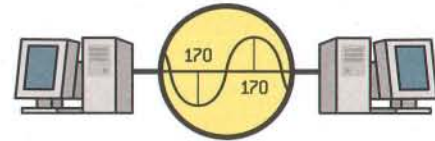
6 Meanwhile Jimmy and Bobby duel to the death in a network video game.

5 ... and plays music streaming from MP3 files on her personal computer upstairs.



DEALING WITH NOISE: ADAPT AND EVADE

POWER LINES inside a house are designed to carry a current that alternates more or less smoothly between -170 volts and $+170$ volts (averaging 120 volts). Communications signals—small voltage variations added to that sine wave—can get clobbered by numerous sources of noise and interference.



IDEAL AC POWER VOLTAGE

Motor Noise

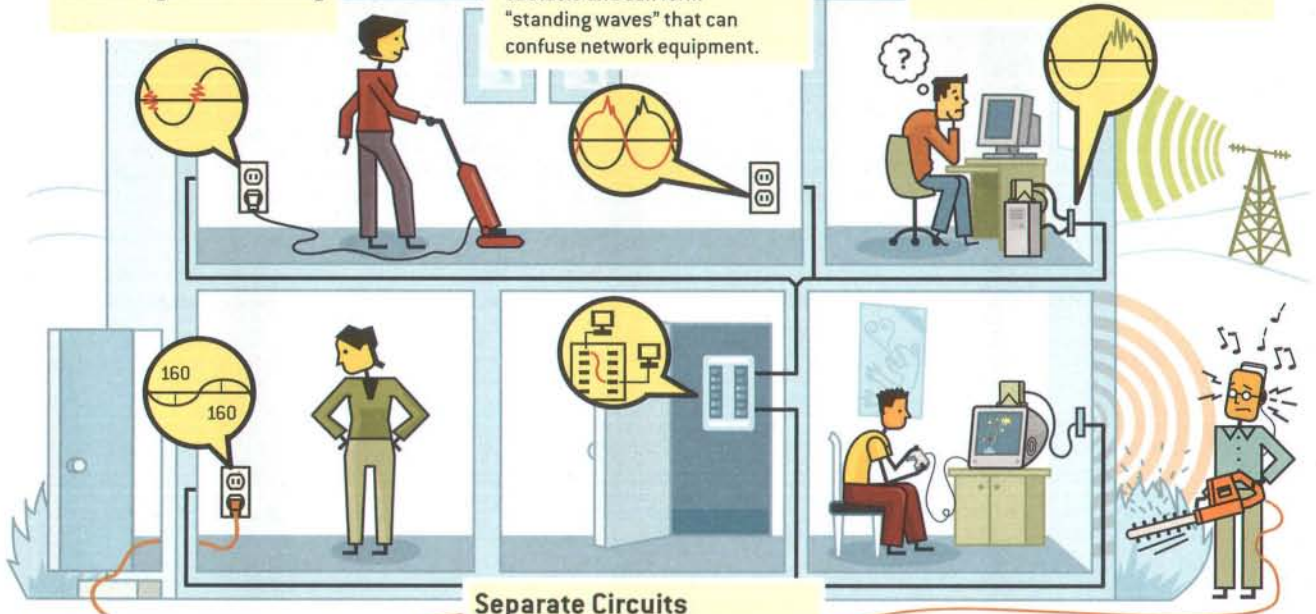
Vacuum cleaners, blenders and other motors sweep brushes across magnets, producing static as a side effect. The static is conducted back through the house wiring.

Reflection

Sudden changes in the resistance of a circuit, such as a termination at an unused outlet, act like a wall to communications signals. They bounce off the obstacle and can form "standing waves" that can confuse network equipment.

Radio Interference

The electrical wiring in a house acts like a large antenna that can pick up signals from radio and television towers or amateur radio operators nearby.



Changes in Size and Shape

Turning on a light switch or plugging in an extension cord changes the size and character of the electrical network, causing brief dips or surges in power levels.

Separate Circuits

Home wiring is split into circuits that connect only at the main breaker or at a transformer down the street. Signals between machines on different circuits must jump those hurdles.

Stray Transmissions

One way to cope with noise is to increase the signal power. But doing so increases stray radio noise transmitted by the antennalike wiring.

How a HomePlug System Avoids the Noise

SMALL SYMBOLS

To send a message, the computer first divides the message into packets.



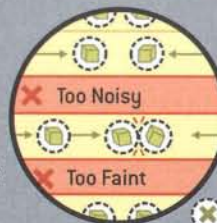
A HomePlug adapter plugged into the wall receives the packets and subdivides each one into "symbols" small enough to fit between bursts of noise.

ERROR PREVENTION CODES

A mathematical algorithm tacks onto the symbol some numbers that describe its contents.



The adapter also pads the symbol with a guard interval that protects the data from collisions with reflected echoes of itself and other symbols.



WIDEBAND MULTIPLEXING

Each pair of adapters scans 76 channels spread across a wide band of spectrum from 4.5 to 21 megahertz. They block those that are too noisy or faint. Symbols are sent simultaneously on all clear channels, but at low power to minimize stray radio transmissions.

DAMAGE REPAIR

The receiving adapter uses another mathematical algorithm to check the integrity of incoming symbols and to repair those with minor damage.



The adapter then reassembles the symbols into a packet, and the target computer stitches the packets into a copy of the original message.

“What makes it all work is a complexity of signal processing that considerably exceeds anything ever used before.” —Oleg Logvinov, president of Enikia

ment. They blocked off the eight channels that would be most likely to interfere with amateur radio operators to head off any complaints to the Federal Communications Commission.

“But we spent most of our effort trying to deal with noise on the power line,” Yonge says. “That was the hardest problem.” The engineers needed a transmission technique that could adapt to rapidly changing conditions on the line. They found one in so-called orthogonal frequency division multiplexing (OFDM), which is used in Europe for digital television broadcasts.

Bending with the Wind

OFDM IS SIMPLER than its intimidating name suggests [see illustration on opposite page]. Unlike FM radio, in which each channel carries a different program, OFDM allows a HomePlug device to send a data message using all 76 independent channels at once. Before two HomePlug adapters start exchanging data, they send each other test signals on every channel and block off those that are too noisy or faint. “That channel map is updated every few seconds,” Yonge explains. So the speed of communication rises and falls as needed for as long as the conversation continues.

That takes care of the wind, to continue our analogy, but it cannot completely compensate for echoes and sudden outbursts of noise. To get around the echoes, each packet of data put on the network is padded with a guard interval, a brief pause that allows the reverberations to die down. And every digital payload is preprocessed to add extra error-checking information. If the message is corrupted en route, the receiver can use the added numbers to mathematically reconstruct the damaged bits.

“What makes it all work is a complexity of signal processing that considerably exceeds anything ever used before in mainstream wireless or wired communications,” says Oleg Logvinov, president of Enikia, a power-line communications company in Warren, N.J. “But the number of logic gates you can pack into a chip today is enough to conquer complexity.” Indeed, Intelon’s chip is nearly as complex as a first-generation Pentium processor but is custom-designed to handle just this one task.

In simulations, Yonge recalls, the system worked great. But the real test came last spring, when Intellon sent engineers out to test its prototype equipment in more than 25 companies and 500 homes of various sizes and vintages around the world. The consortium members had set a high criterion for success, says William E. Earnshaw, Intellon’s head of engineering. “They wanted it to work at full capacity in 80 percent of the homes and at two thirds of full speed in the rest.” The technology passed easily, Earnshaw reports. “You can plug a vacuum cleaner in right next to the modem, and it will still work,” Yonge boasts.

In November, Linksys and Phonex Broadband became the

first to announce retail devices that will work with PCs, printers and any other peripherals that support a USB network connection. Other manufacturers are not far behind. According to Linksys, prices will be similar to those for Wi-Fi gear but a bit higher than HomePNA. “Most every room has at least two outlets,” notes Karen Sohl of Linksys, “so the power-line adapters are more convenient than phone-line products” and thus may command a slightly higher price. The cost of the equipment will fall over time, says Tom Reed, president of HomePlug and manager of emerging technologies at Radio Shack. Because power-line adapters do not need as much analog circuitry, they will always be cheaper to make than wireless devices.

Power-line communications also have a security advantage over wireless, Reed points out. The signals will propagate to other homes on the same transformer—typically three to six houses in the U.S. (or all the units in an apartment building). But the chips automatically come with encryption turned on, unlike Wi-Fi. Cracking the 56-bit ciphers isn’t impossible, but it isn’t easy.

Scherf of Parks Associates perceives two critical tests ahead for the technology. The first is a competing standard being developed by the Consumer Electronics Association. That effort appears to be at least a year away from completion. By the time it arrives, Yonge says, HomePlug will be rolling out version 2, which will increase speeds to 100 megabits a second. But two incompatible standards could sow confusion, nonetheless.

Second, and probably not coincidentally, the companies that make audio, video and telephone components remain on the sidelines. That may delay some of the most appealing applications for power-line communications. “You could plug a cable or satellite video receiver into any outlet in the home and then plug your TV in anywhere and move it around at will,” Reed muses. “Or you could add a virtual second phone line. Say you grew up in Omaha: with voice-over IP [Internet telephony] you could get a local phone number in Omaha that connects over the Internet to ring an IP phone plugged into the wall in your house in California. And when you move to Boston, the number stays the same.”

It will take clever business innovations to make such dreams profitable. But thanks to the innovations of clever engineers, the technology now seems plausible. ■

MORE TO EXPLORE

HomePlug Standard Brings Networking to the Home. Steve Gardner et al. in *Communication Systems Design*, Vol. 6, No. 12; December 2000. Available online at www.csdmag.com/main/2000/12

HomePlug Powerline Alliance: www.homeplug.org

HomePNA: www.homepna.org

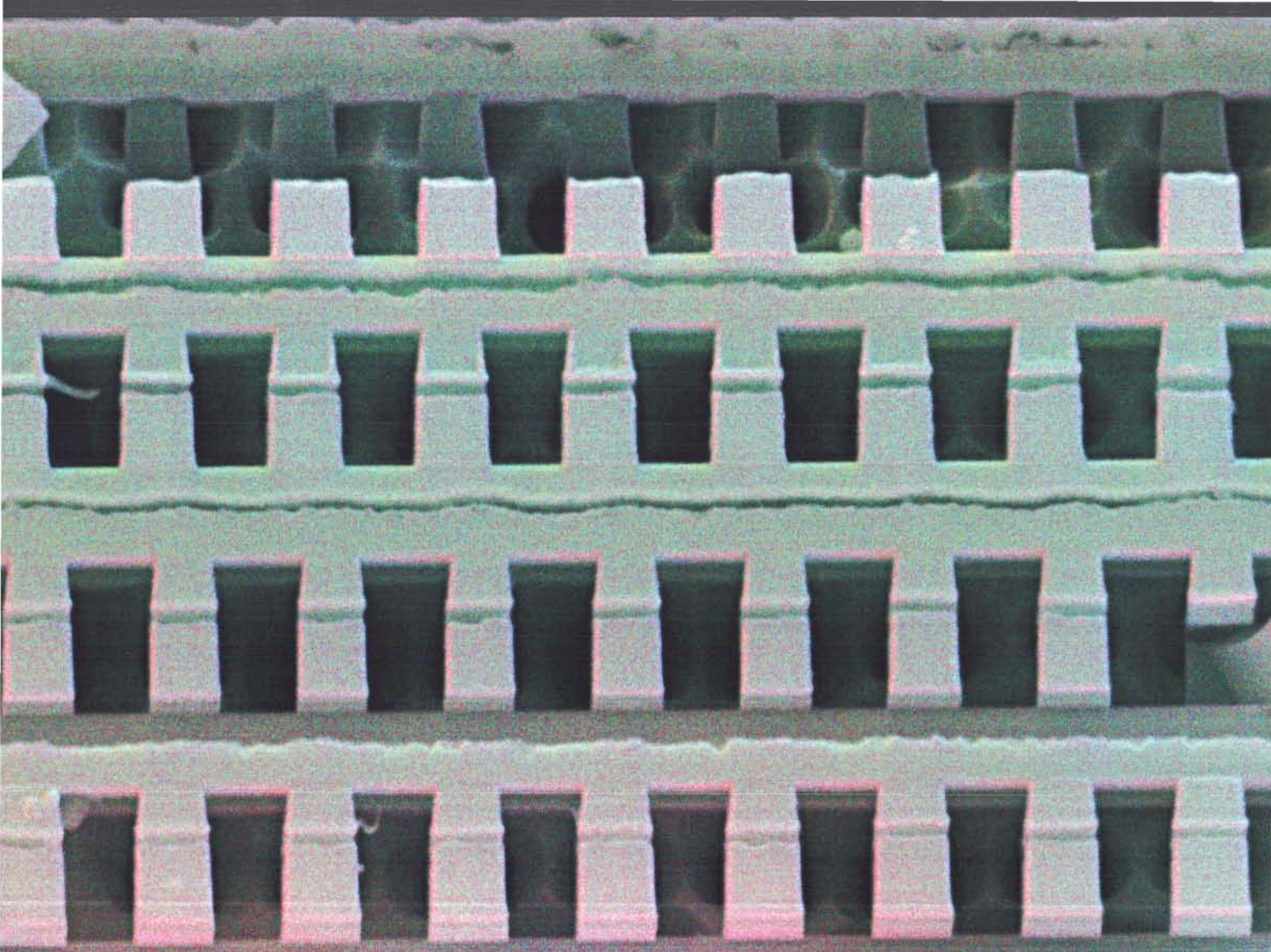
Wi-Fi: www.wi-fi.org

a Vertical

Leap for Microchips

Engineers have discovered a way to pack more computing power into microcircuits: build them vertically as well as horizontally

By Thomas H. Lee



The city of San Francisco stretches over 45 square miles—about twice the area of the island of Manhattan. Yet the economic output of Manhattan dwarfs that of San Francisco. A principal reason for the disparity is that offices in earthquake-prone California tend to spread their workers and machines close to ground level, whereas businesses in New York are stacked vertically into the skies. By building upward rather than outward, developers increase not only the value of their real estate but also the working power of the city as a whole.

An analogous strategy applied to the microscopic world of computer chips could rejuvenate a semiconductor industry that has recently begun to show signs of senescence. Surprisingly, of the more than 100 quadrillion transistors that Intel co-founder Gordon E. Moore estimates have been produced to date, nearly every one has been built on the “ground level,” directly on the surface of silicon crystals. Engineers have accomplished a fantastically regular doubling of transistor density per microchip—we call it Moore’s Law in the industry—simply by expanding the area of each chip and shrinking the size of each transistor. This is like building only shopping malls and no skyscrapers.

VERTICAL STACK of memory cells can store eight bits of information in the area usually allotted to just one bit. Such “3-D” microchips are set to dramatically reduce the cost of digital memory.

That is about to change. For one thing, physicists tell us that Moore’s Law will end when the gates that control the flow of information inside a chip become as small as the wavelength of an electron (on the order of 10 nanometers in silicon), because transistors then cease to “transist.” And many intimidating technical obstacles loom between the current state of the art and that fundamental limit. The trajectory of progress has already begun to droop.

Fortunately, I and other engineers have recently found a way to skirt some of those obstacles, to give Moore’s Law a new breath of life and even to accelerate the delivery of more computing power for less cost. We have shown that it is feasible to make chips that contain vertical microcircuits using the same semiconductor foundries, the same standard materials and similar techniques to those used to manufacture conventional computer chips.

Such “three-dimensional” chips are now being commercialized by Matrix Semiconductor, a company I co-founded in

1998 in Santa Clara, Calif., with computer scientist P. Michael Farmwald and chip design expert Mark C. Johnson. Sometime in the first half of 2002, 3-D memory circuits will hit the market. They will be just the first of a new generation of dense, inexpensive chips that promise to make digital recording media both cheap and convenient enough to replace photographic film and audiotape. In laboratories at Stanford University and Matrix, we have also created prototype devices that incorporate vertical logic circuits. There

tures and can put a billion transistors on a chip. Further improvement should push that limit to 65 nanometers and perhaps 16 billion transistors.

The road beyond that point may be rough, however. Extreme ultraviolet lithography systems that use even shorter wavelengths are just now beginning to function in the laboratory. They still pose many significant problems [see "Getting More from Moore's," by Gary Stix; SCIENTIFIC AMERICAN, April 2001].

This year **3-D memory circuits** will hit the market, just the first of a new generation of dense, inexpensive chips that promise to replace photographic **film** and **audiotape**.

seems to be good reason to expect that even for microprocessors, the sky is the limit.

The Fences of Flatland

TODAY'S STATE-OF-THE-ART microcircuits are not entirely two-dimensional. Intel's Pentium 4 processor, for example, boasts seven layers of wiring, embedded within patterns of insulating material. It is only on the bottom layer of pure silicon, however, that the active semiconducting regions lie.

So far the industry has managed to sustain Moore's Law largely by improving the way it uses that silicon wafer. Materials scientists have invented ways to grow giant crystals of silicon 30 centimeters in diameter that contain less than one part per billion of impurities. Clean-room robots shoot carefully metered doses of ions into wafers cut from these crystals. A process called photolithography defines the ion-activated regions with patterns of light and acid etching to make transistors [see illustration on opposite page]. To cram more transistors onto one wafer requires light of ever shorter wavelength. Mercury vapor lamps have been replaced by deep-ultraviolet excimer lasers that inscribe 130-nanometer fea-

If history is any guide, engineers will probably clear these hurdles; the economic incentive to do so is huge. But as the number of obstacles increases, the pace of progress may slow considerably. The official "road map" published by the Semiconductor Industry Association projects that chips will grow 4 to 5 percent a year in area; historically, area has grown about 15 percent a year. The periodic 30 percent reduction in minimum feature size is probably now going to occur every three years instead of every two. Even at this slower pace, Moore's Law will most likely hit fundamental limits sometime between 2010 and 2020.

One important factor has remained roughly constant: the cost of semiconductor real estate, at about \$1 billion per acre of processed silicon. So why haven't silicon developers taken the seemingly obvious step of building upward? The simplest reason is that transistors are fastest and most reliable only when formed from the perfectly aligned atoms of a wafer cut from a single crystal of silicon.

Once we coat that semiconducting wafer with an insulating oxide or metal wires, there is no known way to recover the underlying crystalline pattern—it's like trying to match the pattern of a parquet floor after it has been covered with carpet. Silicon deposited onto a noncrystalline surface tends to be completely disordered and amorphous. With appropriate heat treatment, we can encourage the silicon to form minuscule islands ("grains") of single crystals, but the ordered lines of atoms collide abruptly at odd angles at the boundaries between grains. Contaminants can pile up at these barriers and short out any transistor or memory cell caught in the middle. For many years, such amorphous and polysilicon (short for polycrystalline silicon) devices were so poor that no one seriously considered them for anything more sophisticated than solar cells.

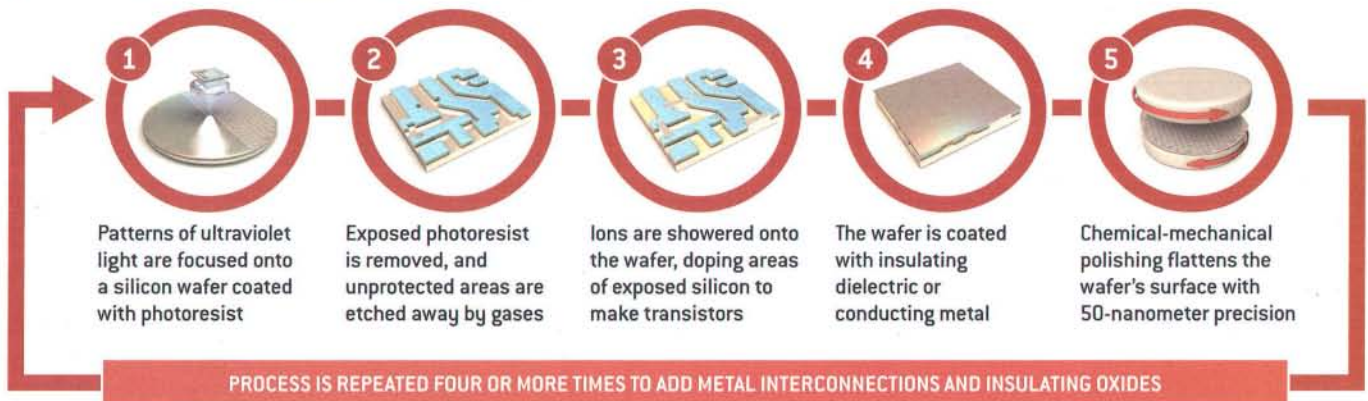
In the early 1980s, however, premature worries that Moore's Law was about to fail stimulated a flurry of attempts to make 3-D microcircuits in which the transistors spanned vertical towers—rather than horizontal bridges—of silicon. James F. Gibbons and others at Stanford used laser beams to

Overview/3-D Microcircuits

- Moore's Law—the steady growth in silicon-based microchip complexity on which the information technology industry depends—is approaching fundamental physical limits. Switching from silicon to new kinds of semiconductors would be enormously expensive.
- Engineers recently have found a way to extend and perhaps accelerate Moore's Law significantly. They have designed and mass-produced multilayered chips in which the semiconducting parts of circuits are no longer confined to a single plane but extend vertically as well.
- The first products incorporating such 3-D microchips—memory cards cheap enough to use as digital film and audio-recording media—are scheduled to appear later this year.

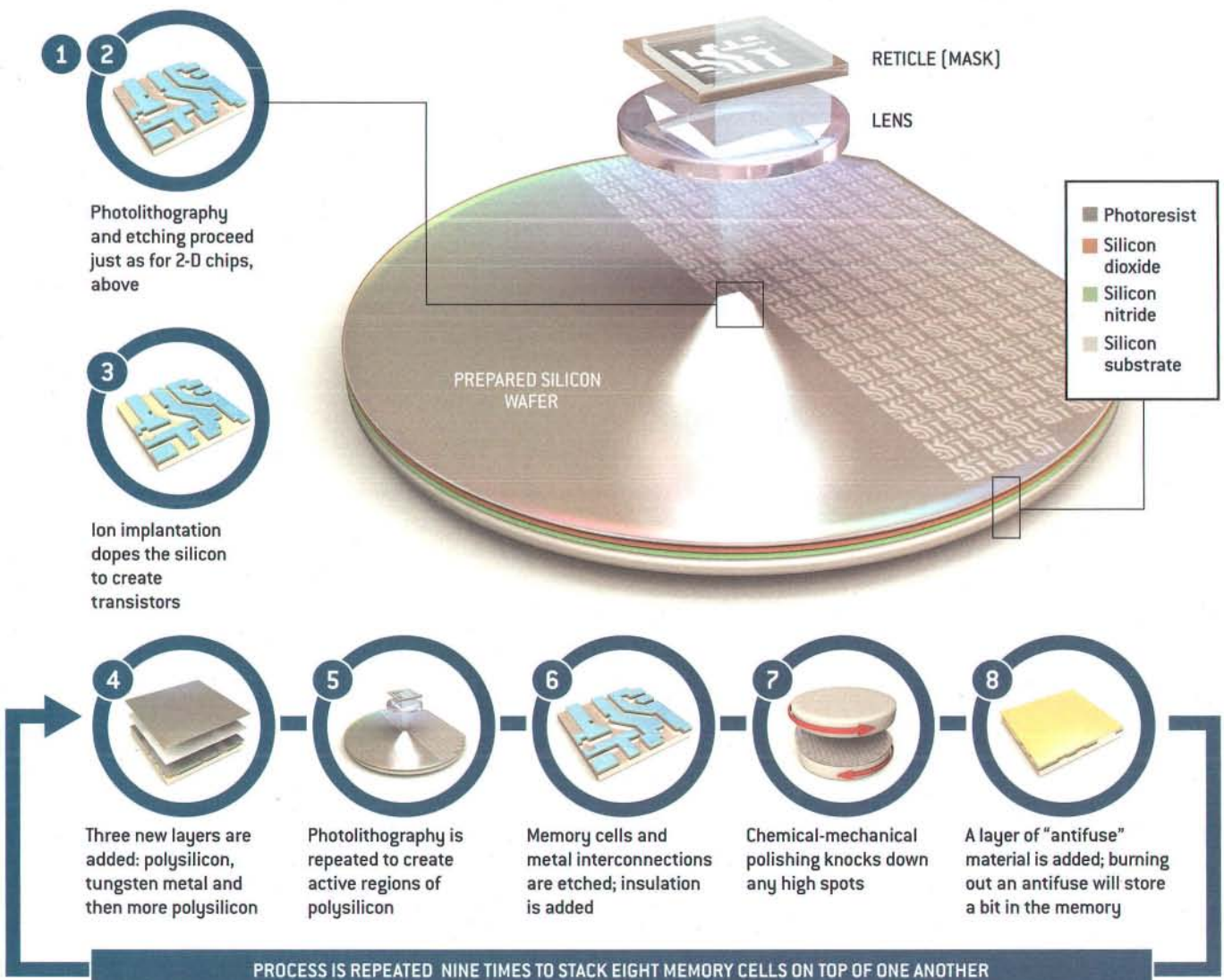
GOING 3-D: SAME EQUIPMENT, DIFFERENT STEPS

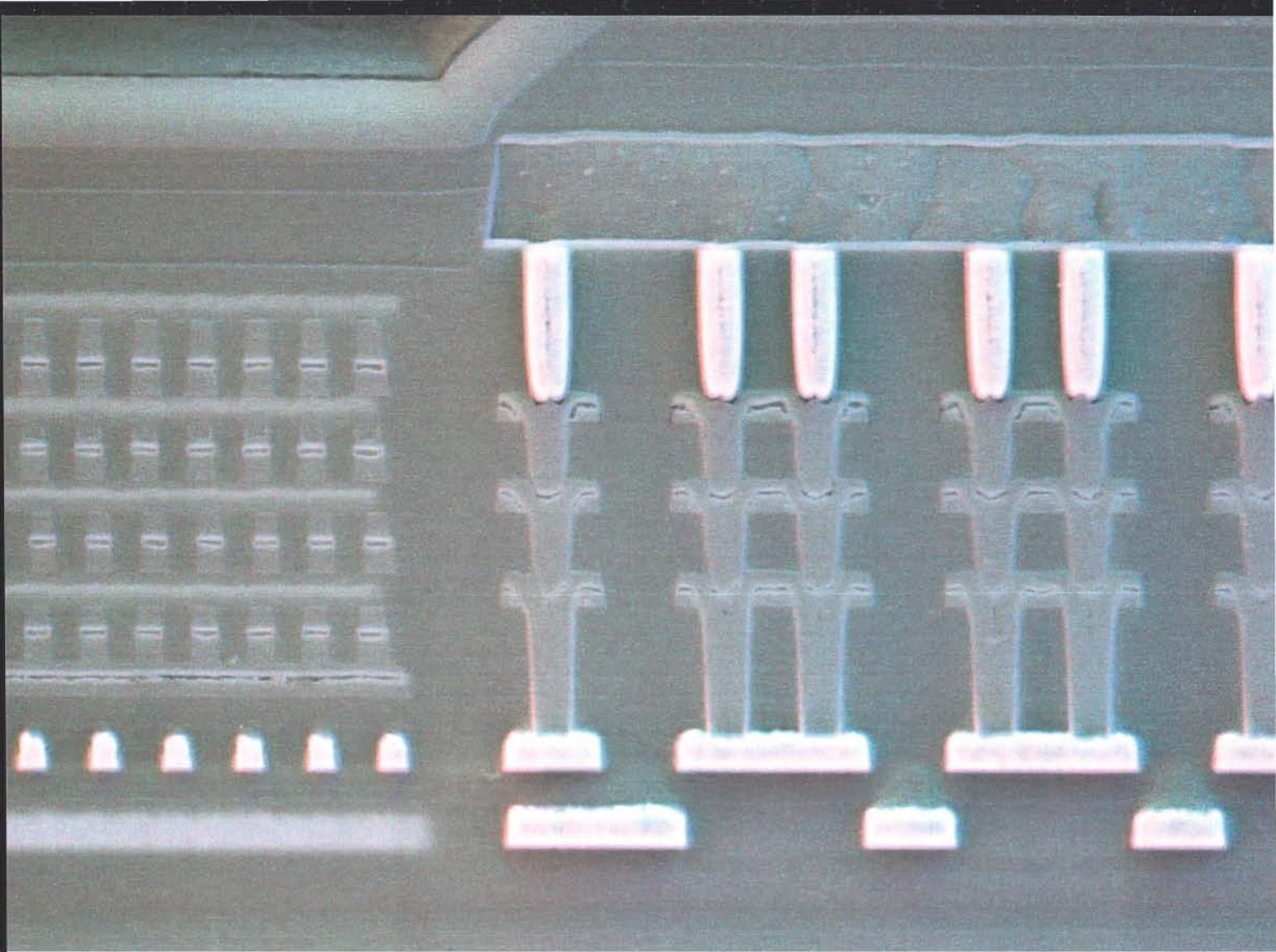
HOW 2-D CHIPS ARE MADE



HOW 3-D CHIPS ARE MADE

Additional steps using the same equipment and materials create 3-D microcircuits with many polysilicon transistors stacked vertically





improve the quality of silicon films deposited onto nonsilicon substrates. Others tried stacking conventional 2-D chips on top of one another. Regrettably, the former approach was too slow and the latter was too expensive to be economically competitive. Traditional chipmaking stayed on track, and engineers stopped thinking much about vertical circuits.

A New Use for Old Tools

IN 1997 FARMWALD and I started exploring 3-D chips again and realized that two key enabling technologies, developed for other purposes, made 3-D circuits truly practical for the first time. One was a technique to lay down polysilicon so that each island of a single crystal is large enough to encompass many memory cells or transistors. The second advance was a way to flatten each coat of new material so that the chips don't rise unevenly like towers built by drunken bricklayers.

We can thank the flat-panel-display industry for the first breakthrough. Its engineers figured out how to make millions of transistors from a thin film spread over a large, amorphous substrate (glass, in their case; other materials in ours). Thin-film transistors now populate the display panels of virtually every laptop computer. Part of the secret is to deposit the silicon at about 400 degrees Celsius as an extremely

LAYERS OF POLYSILICON that form the honeycomb of memory cells (*left*) are interconnected by "vias" (*vertical columns at right*). These are connected in turn by tungsten wires (*bright structures*).

smooth (though amorphous) film, then to cook the entire wafer uniformly above about 500 degrees C for a few minutes. This converts the film to polysilicon with regular crystalline regions of a micron or more in diameter. Although LCD panels require only a single layer of transistors, the same machines that make the panels can also manufacture multilayer devices.

THE AUTHOR

THOMAS H. LEE worked as a circuit designer at Analog Devices in 1990, after receiving his doctorate from the Massachusetts Institute of Technology. In 1992 he joined Rambus to develop high-speed analog circuitry for memory chips. His work was also incorporated into microprocessors made by Advanced Micro Devices and Digital Equipment Corporation. In 1994 Lee joined the electrical engineering faculty at Stanford University, where his research focus has been on gigahertz communication circuits. He is a Distinguished Lecturer of two professional societies and a Packard Fellow and has been granted 14 U.S. patents, including four for constructing and mass-producing 3-D integrated circuits.

The second key enabling advance, called chemical-mechanical polishing (or CMP), emerged from IBM's research labs in the late 1980s. Back then, chip designers considered it risky to add two or three layers of metal on top of the silicon wafer because each new layer added hills and valleys that made it difficult to keep photolithographic patterns in focus.

To eliminate the bumps in each layer, process technologists adapted a trick that lens makers use to polish mirrors. The basic technique was used on all Intel 80486 processors: after each coating of silicon, metal or insulating oxide is added, the wafer is placed facedown on a pad. Spindles then rotate the pad and wafer in opposite directions while a slurry of abrasives and reactive alkaline chemicals passes in between. After mere minutes of polishing, the wafer is flat to within 50 nanometers, an ideal substrate for further processing. With advances in CMP machines, seven and eight layers of metal have become common in microchip designs; patience seems to be the main limiting factor in adding still more layers.

Building directly on these 2-D technologies, we have made 3-D circuits by coating standard silicon wafers with many successive layers of polysilicon (as well as insulating and metallic layers), polishing the surface flat after each step. Although electrons do not move quite as easily in polysilicon as they do in the single-crystal kind, research has produced 3-D transistors with 90 to 95 percent of the electron mobility seen in their 2-D counterparts.

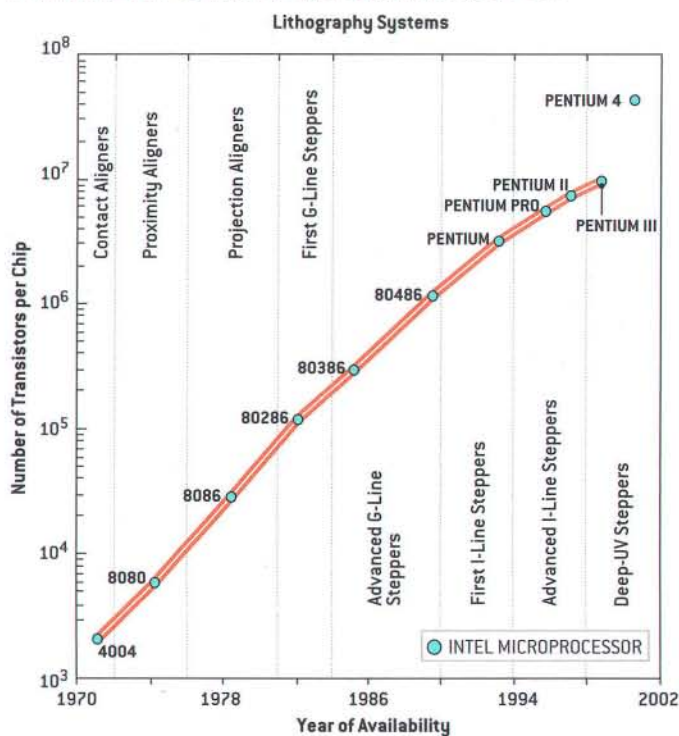
Stacking devices vertically offers a way around some of the weighty obstacles that threaten to derail Moore's Law. As shopping-mall-style chips continue to sprawl outward, for example, it becomes increasingly hard to keep the photolithographic image in focus at the edges. And the relatively long wires that connect far-flung sections of conventional microprocessors cause delays that reduce performance and complicate design.

Ever shrinking circuits pose other problems. Transistors depend critically on a thin insulating layer below the control electrode. In the most advanced 2-D chips, this layer of silicon dioxide insulation measures just three nanometers—about two dozen atoms—in thickness. From transistor to transistor, that thickness must not vary by more than one or two atoms. The industry routinely meets this challenge, because it is much easier to grow superthin films than it is to etch supernarrow channels. But there may be no practical way to make these insulating layers much thinner, because current flow by quantum tunneling makes them progressively worse insulators. It's likely that some other material will soon have to replace silicon dioxide, but toolmakers have yet to agree on what that material will be.

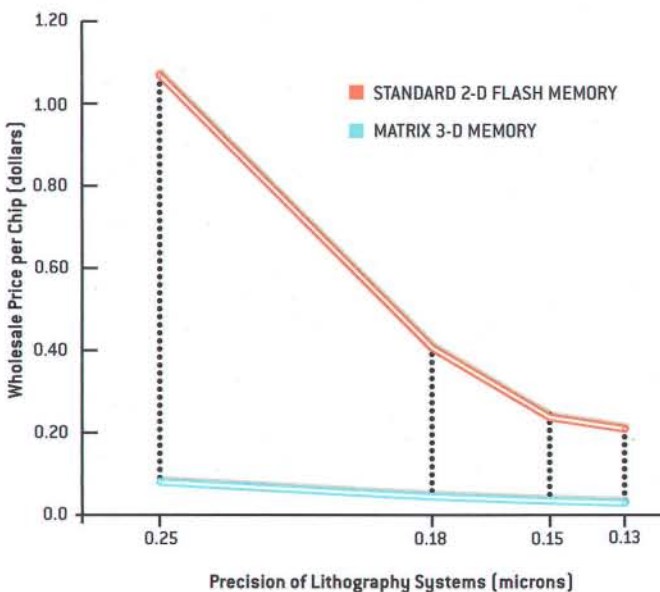
There have been many novel chip designs proposed to address these problems. Most depend on replacing silicon altogether with various exotic materials, such as organic polymers, carbon fullerenes, copper compounds, ferroelectrics or magnetic alloys. But to abandon silicon is to squander an enormously valuable foundation of knowledge constructed

MOORE'S LAW—the periodic increase in microchip complexity—slowed around 1990 as engineers struggled to invent photolithography systems that etch ever smaller structures into silicon (top). Although architectural changes boosted the size and number of gates in the Pentium 4, Moore's Law will likely slow further as lithography systems move from ultraviolet light toward x-rays, which are much harder to focus. By increasing the vertical complexity of microcircuits, however, chip designers can achieve the same computing power at far lower cost (bottom).

EACH NEW ADVANCE IN CHIPMAKING ...



... LOWERS THE COST PER CHIP



Vertical electronics can reduce

manufacturing costs 10-fold or more, and the density of 3-D devices should increase at least as fast as **Moore's Law** as we add layers.

over 50 years with some \$100 billion worth of investment.

The 3-D electronic design process, in contrast, introduces no new atoms and leverages the huge industry investment in thin-film and CMP equipment. Because it is so expensive to produce and process ultrapure silicon ingots, the cost of silicon is largely proportional to the area (not the volume) consumed. So vertical electronics can reduce manufacturing costs 10-fold or more compared with traditional chips. And the density of 3-D devices should increase at least as fast as Moore's Law as we add more and more layers.

Digital Film and Beyond

TRADITIONALLY, semiconductor companies have worked the bugs out of new fabrication processes by making memory devices before attempting to mass-produce more complicated chips such as logic circuits. Memories are vast arrays of fundamentally simple cells, so there are fewer skills to master and fewer problems to solve.

That is the approach we at Matrix will take later this year as we introduce a 3-D memory chip in which the cells are stacked eight high [see illustration below]. Unlike the RAM memories used in PCs, these chips use exceedingly

simple memory cells that make them more like film, indelible once written. They are intended to be a low-cost medium for digital photography and audio. With 512 million memory cells, this first vertical microchip has enough capacity to store more than an hour of high-quality audio (through data compression) and a few hundred photographs (each comprising about one million pixels). The capacity will rise, and the unit cost will fall, over time. We have already proved that 12-cell-high devices are feasible, and 16-layer chips seem well within reach.

We have also demonstrated much more complex 3-D microcircuits in the laboratory, including static RAM, logic gates and even erasable EPROM memories. Although they are in very early stages of development, these basic building blocks are all that is needed to recast any planar circuit—including dynamic RAM, nonvolatile memories, wireless transceivers, and microprocessors—in 3-D form. Stood on end, the transistors in such circuits could be quite tiny because their channels will be made from thin films that are 10 times as precise as channels defined by ultraviolet light.

As with all engineering advances, this new manufacturing technique has limitations and trade-offs. Some fraction of memory cells or transistors in a vertical microcircuit will happen to straddle a boundary between polysilicon grains and will possibly fail as a result. We will have to use error detection and correction routines, like those used with audio CDs, and find ways to route signals around defective paths. The strategies of fault-tolerant computing, though well known, have generally

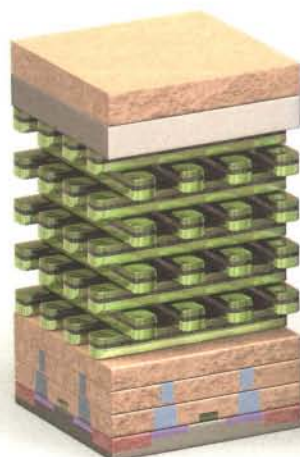
INTERIOR STRUCTURE of 3-D chips is dramatically different from those of conventional, 2-D memories and microprocessors. In standard logic and memory chips, all the transistors in the circuit are confined to a single crystalline layer of silicon. In 3-D microcircuits, transistors or memory cells are formed within multiple layers of silicon:

3-D VOLATILE MEMORY
(MATRIX SEMICONDUCTOR)

2-D RANDOM-ACCESS MEMORY
(IBM 256-MEGABIT)

3-D LOGIC CIRCUIT
(LAB PROTOTYPE)

2-D MICROPROCESSOR
(ADVANCED MICRO DEVICES ATHLON)



■ Monosilicon substrate

■ Insulators

■ Aluminum wires

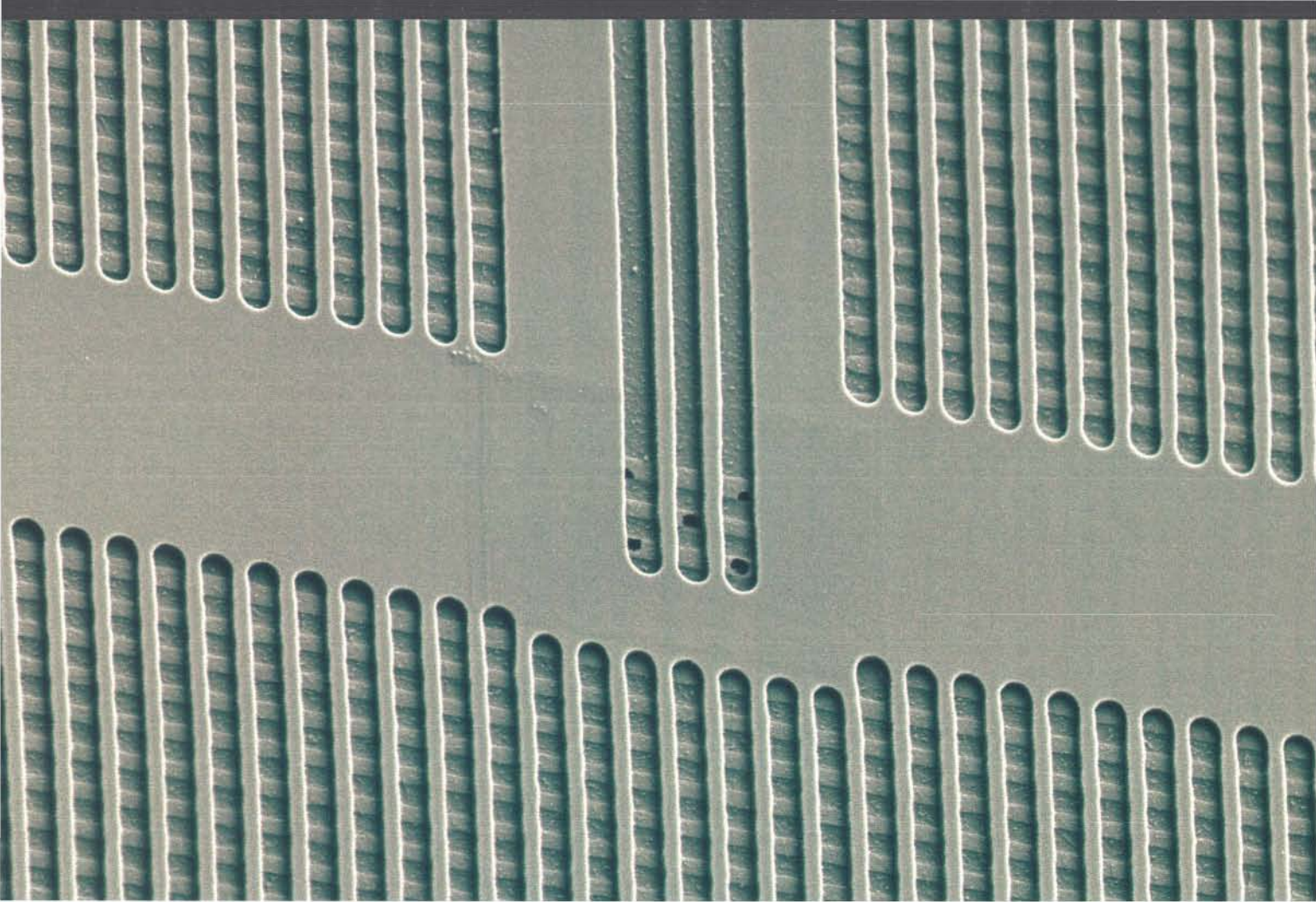
■ Polysilicon

■ Tungsten plugs

■ Ion-doped silicon

■ Isolation oxides

NOT TO SCALE



not been built into microchips themselves. Such techniques are unnecessary and too cumbersome for application in most planar contexts, but the cost reductions afforded by 3-D processing fortuitously make the remedial technology economically feasible precisely when it becomes necessary.

Speed is another trade-off. Modern thin-film transistors typically perform at about half the speed of monocrystalline devices, although the difference is smaller when you compare entire circuits, because components packed in three dimensions need considerably shorter wires. Numerous researchers are investigating ways to close that gap further.

Beyond those special considerations, 3-D chips face essentially the same challenges as do conventional planar electronics—certain problems just appear sooner because of the effective acceleration of Moore's Law. Heat may be the most acute issue for dense 3-D devices because of their smaller surface area. The power density of a modern microprocessor already exceeds that of the burner on a typical stove. Ineffectiveness of current strategies for dissipating all that heat, such as reducing voltages or selectively activating only parts of a circuit, may limit the performance of dense 3-D circuits unless more advanced cooling technology is used. Fortunately, the newest microrefrigerators can now remove 200 watts per square millimeter while consuming only about one watt. Thermal limits are thus not yet fundamental impediments.

There is certainly lots of room for improvement. The fluid-

BITS ARE STORED permanently in this 3-D memory chip when antifuses are blown (dark spots in center), connecting two halves of a circuit.

cooled human brain, whose dimensions considerably exceed those of any 3-D circuit currently contemplated, dissipates a mere 2.5 watts; a 2.2-square-centimeter Pentium 4 microprocessor, in contrast, consumes about 80 watts. Although we cannot rule out the possibility that the inability to solve the heat problem may ultimately impose harsh limits on what 3-D circuits can do, history suggests that the strong economic incentives at play will eventually spark creative solutions.

Enabling Moore's Law to continue even a few years longer than it otherwise would have will have far-reaching consequences. For 30 years, chip manufacturers have striven constantly to print ever smaller structures within a single plane. It seems inevitable that in the future we will scale microcircuits vertically as well as horizontally. The technology is both possible and practical, and the benefits are far too compelling to ignore.

MORE TO EXPLORE

The Solid-State Century. *Scientific American* Special Issue, Vol. 8, No. 1; October 1997.

Multiple Layers of CMOS Integrated Circuits Using Recrystallized Silicon Film. Victor W. C. Chan et al. in *IEEE Electron Device Letters*, Vol. 22, No. 2, pages 77–79; February 2001.

The Electronic

OFFERING A GLIMPSE

of a future with rewritable periodicals, this E Ink Corporation prototype "prints" text using electronic ink. Voltages are supplied to the ink by a thin-film-transistor panel, from IBM. The panel is 800 by 600 pixels; each pixel is formed by charged pigment—the "ink." Electrically erasable programmable memory sticks (*sitting atop display, at right*) are used in setting the text.



Digital "paper" that displays changing text and graphics would ideally marry the best features of traditional printed materials with those of video screens. Companies are racing to realize that promise using two competing technologies. Already retailers are testing cost-saving changeable e-ink signage.

PAPER CHASE

Pliable, updatable e-newspapers, e-books and even an e-SCIENTIFIC AMERICAN could be here within a decade.

BY STEVE DITLEA

i

t offers excellent resolution and high contrast under a wide range of viewing angles, requires no external power to retain its image, weighs little, costs less and is remarkably flexible (literally and figuratively)—unlike today's computer displays. No wonder traditional ink on paper continues to flourish in a digital world that was expected to all but do away with it.

Yet ink on paper is lacking in one of the essential traits of computer displays: instantaneous erasure and reuse, millions of times without wearing out. Electronic ink on paper with this

ability could usher in an era of store signs and billboards that could be updated without pulping acres of trees; of e-books that embody the familiar tactile interface of traditional books; of magazines and newspapers delivered wirelessly to thin, flexible page displays, convenient for reading, whether on crowded subways or desert islands.

There have been intermittent efforts to produce such electronic paper over the past three decades, but only recently has research gone into full swing. The day when *Scientific American* and other periodicals are routinely published in this medium may come before 2010, thanks to competition between two start-up firms. Both companies are offshoots of major research institutions: the Xerox Palo Alto Research Center (PARC) and the Massachusetts Institute of Technology Media Laboratory. Both firms base their core technologies on tiny, electrically charged beads, with the imaging capability controlled electronically. And they are not only racing each other to commercialize their efforts but are also anticipating competition from the organic light-emitting diodes that are beginning to emerge from laboratories.

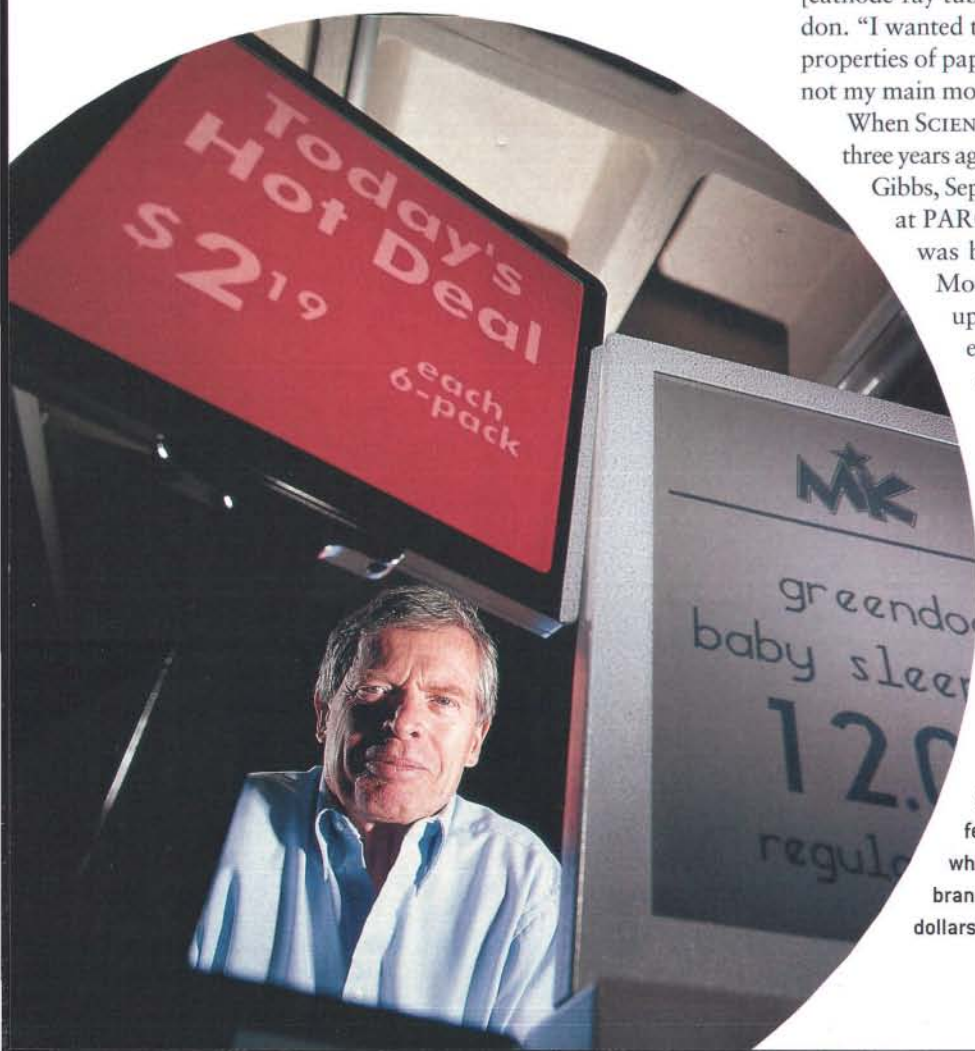
Spinning Off Electric Paper

THE EARLIEST ATTEMPT at "electric paper," as it was originally called, came as a response to the poor visual quality of the computer displays available in the early 1970s. "The CRTs [cathode-ray tubes] were too dim," recalls Nicholas K. Sheridan. "I wanted to find a display material with as many of the properties of paper as possible. Finding a paper substitute was not my main motivation."

When *SCIENTIFIC AMERICAN* last caught up with Sheridan three years ago [see "The Reinvention of Paper," by W. Wayt Gibbs, September 1998], he was a senior research fellow at PARC, demonstrating prototypes of what Xerox was by then calling "electronic reusable paper."

More than 20 years earlier at PARC he had come up with the basic idea for this display medium, embedding plastic beads scarcely the width of a human hair in a flexible transparent film. Each bead is two-toned: one half white and one half black, with an opposing electrical charge on each half. Apply an appropriate electric field to the transparent surface, and a bead can be rotated to lock either a white or black dot onto the viewing plane—creating, in effect, ink that twists itself into the right place.

THREE DECADES AFTER his initial vision of creating an electronic display with as many of the features of paper as possible, Nicholas K. Sheridan of Gyricon Media demonstrates the feasibility of SmartPaper displays. The displays, which are now being marketed under the MaestroSign brand, could save individual stores thousands of dollars in signage costs.



Sheridon called his invention Gyricon, Greek for "rotating image." Soon shelved by Xerox managers who were more interested in exploring new printing technologies than in making displays, the reusable-paper concept wasn't revived until 15 years later—hence Sheridan's demo described in these pages at the time.

He would have to wait still longer for his technology to make its way into a commercial product. First it had to be spun off into a separate company. In December 2000 Gyricon Media was launched as an independent venture headquartered in Palo Alto, Calif. (with Xerox retaining a majority financial stake in the company). By March 2001 the new firm made its first product announcement at the GlobalShop trade show for retail store displays in Chicago. On the floor of the McCormick Place convention center, Sheridan, now Gyricon Media's research director, could be seen admiring a smoky green 11-by-14-inch panel on an aluminum stand, the kind you find on department store tables. This sign cycled through several sales messages and the price "\$89.99" in slightly speckled type, and it would be powered by its three AA batteries for up to two years. Nearly 30 years after its original inspiration, Sheridan's electric paper was finally available—with limitations.

This prototype of what is now being marketed as SmartPaper was to be field-tested throughout the second half of 2001 in 15 signs at the Macy's department store in Bridgewater, N.J. Resolution was the equivalent of a modest 100 dots per inch (dpi). In comparison, the resolution of the print version of this magazine is 1,200 dpi. And because inflexible electrodes were used to activate the pliant SmartPaper material (a silicone rubber film soaked in oil to allow the bichromate beads to rotate), this version of e-paper was rigid as a board.

By 2002 Sheridan expects the commercial sale of similarly sized signs that can be easily updated via a wireless network. To a retail client like Federated Department Stores, Macy's parent company, which is currently spending more than \$250,000 a week on changing its in-store signs, such renewable signage could prove highly desirable. Also due out next year are smaller SmartPaper signs meant to keep prices up-to-date on supermarket shelves, where inaccurate numbers can turn into expensive fines under item-pricing laws.

The pliable, reusable e-newspaper or e-magazine of the future "could happen in a few years," Sheridan has predicted on several occasions. He happens to have a concept model: a slit aluminum cylinder from which he pulls out a sheet of SmartPaper, papyrus scroll-like. In a working model, an array of electrodes along the edge of the cylinder would imprint up-to-the-minute news or feature stories on the paper's flexible, rubbery surface; plastic sheets would protect the paper from being damaged. Smaller-size beads necessary for higher resolution are on the way. As for a full range of colors, Sheridan has been issued



JOSEPH JACOBSON of the M.I.T. Media Lab and E Ink seeks to create the effect of real paper that prints itself. His ultimate goal is a cybercodex, "the last book," a bound volume of hundreds of e-ink pages with enough memory chips to store the entire contents of the Library of Congress.

a patent for subtractive color using transparent Gyricon beads with thin disks of color filter material in cyan, magenta and yellow, each addressable by different voltage levels.

Nevertheless, as paperlike as it may become, this electronic paper may never feel exactly like the original. Sheridan admits, "It will never be as light as paper. Paper is about four mils thick; this will always be 12 or 15 mils thick. But it doesn't have to exactly replicate paper to be useful."

Making a Mark with E-Ink

REAL PAPER ABLE to print itself was the starting point for Gyricon Media's principal rival in the digital paper market. Independent of Sheridan, in 1995 Joseph Jacobson, then a Stanford University physics postdoctoral researcher, was looking for an interesting problem to tackle. He came up with the notion of a book full of pages that could be electronically reconfigured to display the text of *King Lear* or *General Relativity* or any of hundreds of other tomes stored in silicon memory in the book's spine.

For his imaging technology, Jacobson turned to electrophoresis, the movement imparted by an electric field to charged

particles that are suspended in a liquid. In place of pigment-carrying beads, he used transparent polymer microcapsules containing a blue liquid dye and white particles. When the positively charged particles of white titanium dioxide remain on the viewable side of the microcapsules, they produce a white page; a negative charge on an electrode below a capsule will draw these particles to the other side, creating an inklike image in their place—until an opposite electrical pulse sends the white pigment back. Reversing this process produces white letters on a dark background. Suspended in water, the microcapsules can be printed on paper or electrode-bearing materials just like ink. Jacobson called this “electrophoretic ink,” or e-ink.

Appointed to an assistant professor position at the M.I.T. Media Lab in 1995, he continued his research into e-ink with two of his undergraduate students, J. D. Albert and Barrett Comiskey. In 1997 the three of them, along with Harvard Business School graduate Russell J. Wilcox, founded E Ink Corporation in Cambridge, Mass. The start-up soon attracted funds from venture-capital firms, corporate investors, including Motorola and the Hearst Corporation, and an R&D grant from the Defense Advanced Research Projects Agency (DARPA).

In 1999 E Ink introduced the first store signs using its technology. The rigid signs, released under the brand name Immedia, measured six by four feet and displayed large, white italic characters (with resolution equivalent to 2 dpi) on a blue background. They were tested in several J. C. Penney stores, as well as by the *Arizona Republic* newspaper for headline displays, and even in sandwich boards worn on city streets for Yahoo. Research found the signs effective in increasing store traffic and departmental sales for E Ink’s customers in retailing, but customers wanted a greater choice of fonts, colors and graphics. Until E Ink can develop its next generation of store signage, it has withdrawn from this market.

Jacobson is not involved in E Ink’s day-to-day operations (he serves on its board of directors, while devoting most of his time

to his position as director of the Molecular Machines Research Group at the Media Lab). In his stead, Michael D. McCreary is E Ink’s vice president of research and development. At the firm’s offices in an office park on the outskirts of Cambridge, he explains that in the company’s business plan, store signs were always considered a first step in proving the viability of electronic ink. “Our next step is developing high-resolution displays for portable devices,” McCreary says. He shows a rigid, Palm Pilot-style screen (with better contrast, at 80-dpi resolution) that’s viewable at wider angles than the standard black-and-white liquid-crystal display (LCD) on a handheld computer.

In February, Philips Components, a division of Royal Philips Electronics in the Netherlands, secured exclusive global rights for a period of time to manufacture and sell display modules using E Ink technology for personal digital assistants (PDAs) and electronic books in exchange for a \$7.5-million investment in the company. With this version of electronic paper drawing as little as one hundredth the power of a comparable LCD screen, Philips is banking on a competitive advantage in battery life for its handheld devices, which will be available within the next two years.

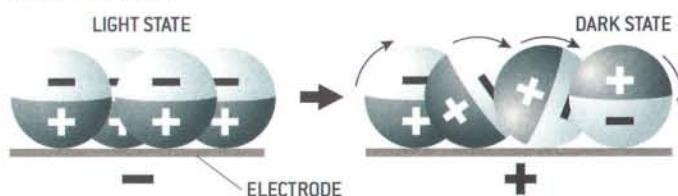
Another corporate partner, Lucent Technologies, licensed its Bell Labs-developed plastic transistor technology to E Ink, while also investing in the company. In November 2000 this alliance demonstrated the first flexible electronic-ink display—a five-by-five-inch screen with the consistency and thickness of a computer mouse pad and just 256 cornflake-size pixels, which cycled through checkerboard patterns, the companies’ names and the E Ink corporate logo. This demo was a proof-of-concept that the active-matrix circuitry for addressing E Ink’s microcapsules could be fabricated with plastic materials rubber-stamped onto a flexible sheet of plastic.

By April, E Ink and yet another partner, IBM Research, had announced their first higher-resolution active-matrix electronic-ink display—a 12.1-inch-diagonal screen with a resolution comparable to 83 dpi, or about the same sharpness as a typi-

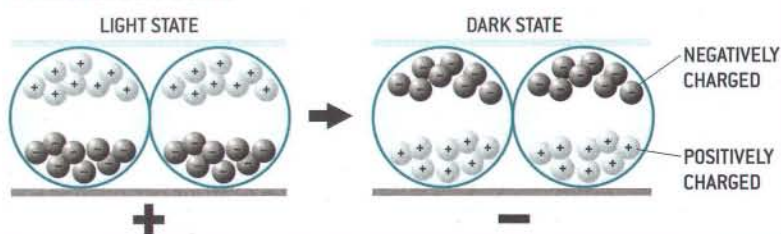
HOW E-PAPER WORKS

Both technologies being developed commercially for electronically configurable paperlike displays rely on microscopic beads that change color in response to the charges on nearby electrodes. Gyricon Media’s SmartPaper uses two-tone solid beads that twist around in place (*top*). Inventor Nicholas K. Sheridan’s breakthrough for producing tiny symmetrical beads involved pouring black and white resins onto a rapidly spinning disk. E Ink’s Electronic Ink employs see-through microcapsules containing pigment chips that move through a liquid medium (*bottom*). Manufacturing applies standard techniques developed for microencapsulated coatings on business forms.

GYRICON BEADS



E INK MICROCAPSULES



cal laptop computer screen. To match the requirement of IBM's electronics, E Ink researchers made their microcapsules change color 10 times faster than they did in the original formulation. For better contrast, the encapsulated dye's color was changed from blue to deep black.

In May, E Ink and Japan's Toppan Printing Company introduced a prototype color electronic-ink display. Using Toppan's color filter arrays, which are now widely deployed in standard LCDs, the demonstration screen showed eight colors. Using this technology, E Ink expects to produce displays capable of showing 4,096 colors, comparable to handheld computer and game screens.

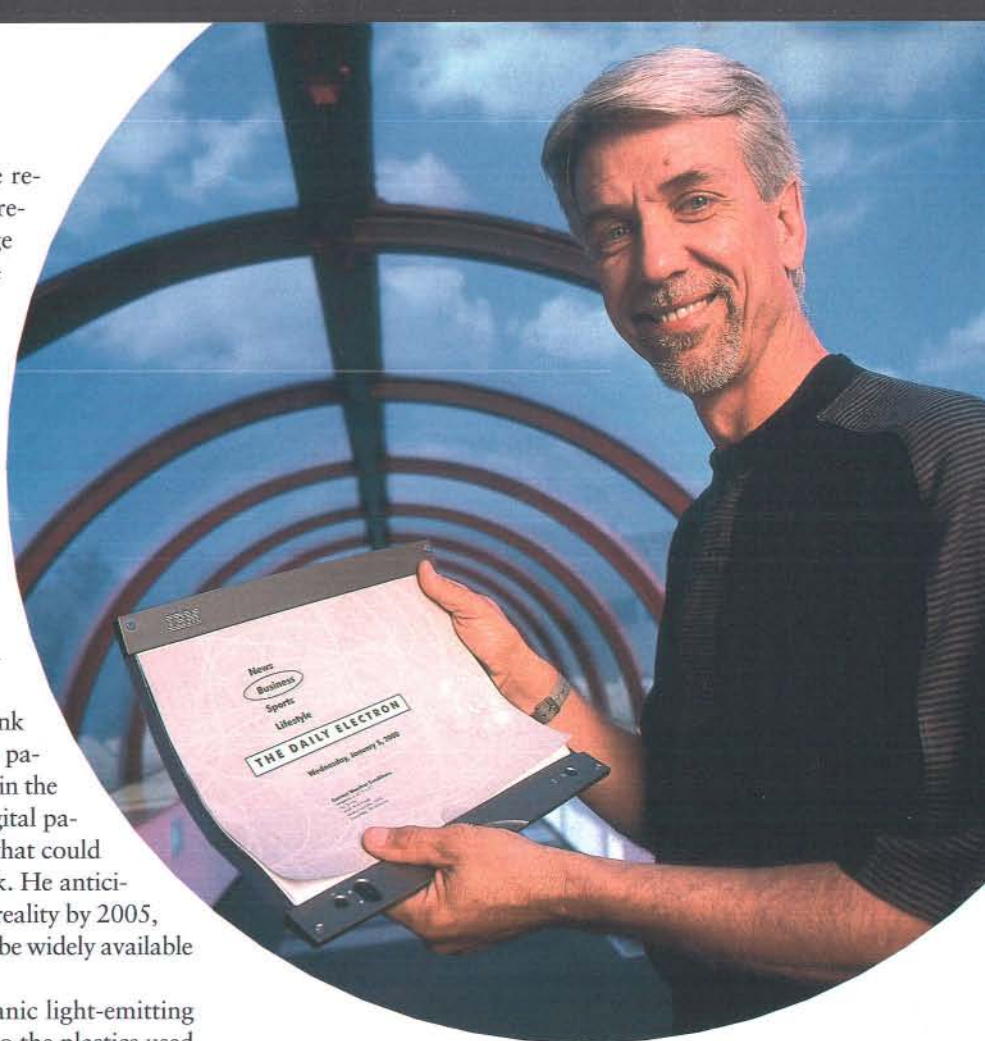
These recent prototypes have brought E Ink closer to its ultimate goal. "We call it 'radio paper,'" McCreary explains of the third stage in the E Ink business plan. This will be flexible digital paper with high-resolution-color capabilities that could be reconfigured via a wireless data network. He anticipates that radio paper will be a commercial reality by 2005, at which point similar technologies may also be widely available from Gyricon and other sources.

E Ink will also be competing with organic light-emitting diodes. Carbon-based compounds similar to the plastics used in E Ink and Lucent's flexible display can produce light-emitting semiconductors that are also pliable and relatively energy-efficient. That this alternative to electronic paper is being developed by Eastman Kodak, IBM and other well-financed firms should soon make this technology a credible challenger.

The Last Book

ALMOST FROM THE BEGINNING, Jacobson's long-term vision for e-ink has included "the last book": several hundred bound pages of self-printing paper with a separate processor imprinted on each page and enough memory chips in the hard-cover volume's spine to store the entire contents of the Library of Congress. With a single page of e-inked paper able to replicate any stored page of text, graphics or even video, why bother binding together so many pages? According to Jacobson, one reason is to engage a reader's spatial memory: thumbing through a book-length work makes it easier to locate a particular passage or illustration.

Somewhere between Jacobson's e-tome and Sheridan's e-scroll, there's another format that electronic paper publishing could adopt. This one is an updated variation on early printing's folios—binary multiples (8, 16 or 32) of pages cut from a single large printed sheet. In 1999 Robert Steinbugler, head of IBM's corporate strategic design program, unveiled a design prototype for the eNewspaper—a rubberized, flexible, portfolio-style display device containing eight two-sided pages made of digital paper (actually plastic mock-ups, for now). Based on



ROBERT STEINBUGLER, head of IBM's corporate strategic design program, shows a rubberized, flexible portfolio-style display concept with eight two-sided pages made of digital paper. The eNewspaper combines the familiar experience of flipping pages with the convenience of instantly rewritable text.

interviews with newspaper publishers and readers, Steinbugler concluded that a sheaf of pages afforded the ability to flip back and forth among stories without having to redraw their text, while also offering the serendipitous juxtaposition of stories that still distinguishes newspapers in print from their online, one-screen-at-a-time versions.

Given today's accelerated quest for electronic paper, it may not be too long before *Scientific American* readers are offered a choice of e-folio, e-hardcover or e-papyrus editions. ■

Steve Ditlea is a freelance journalist based in Spuyten Duyvil, N.Y. He has been covering technology since 1978.

MORE TO EXPLORE

Information about Electronic Reusable Paper is available on the Xerox PARC Web site at www.parc.xerox.com/dhl/projects/gyricon/

Information about SmartPaper is available on the Gyricon Media Web site at www.gyriconmedia.com/smartpaper/index.asp

What Is Electronic Ink? Available on the E Ink Web site at www.eink.com/technology/index.htm

The Last Book. Joseph Jacobson in *IBM Systems Journal*, Vol. 36, No. 3; 1997. Available at www.research.ibm.com/journal/sj/363/jacobson.html

Code Red

for the web

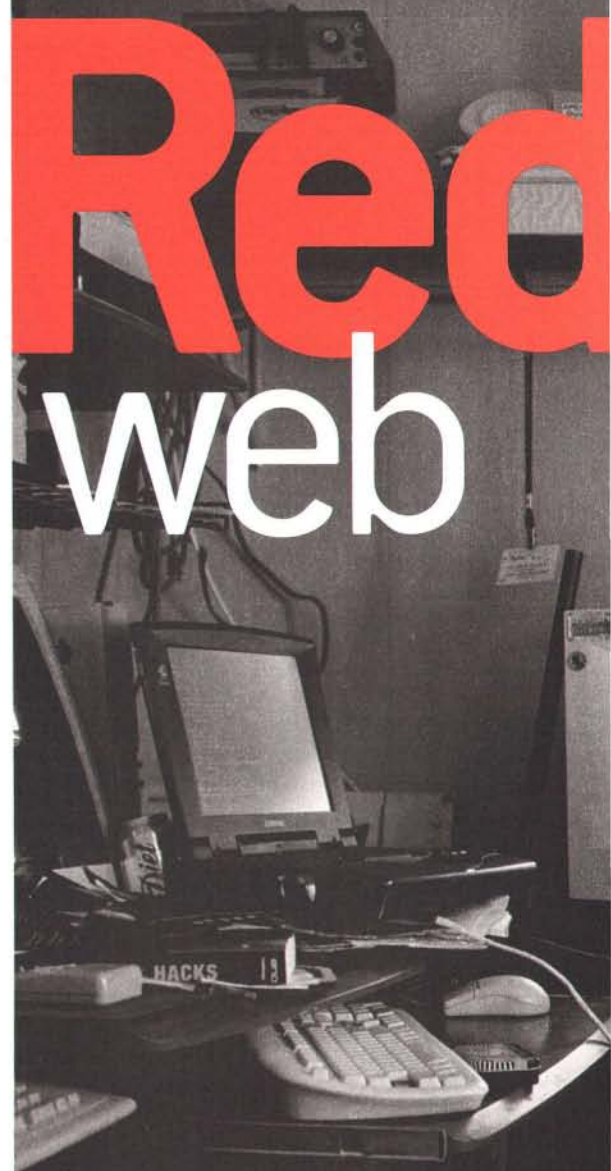
“Imagine a cold that kills. It spreads rapidly and indiscriminately through droplets in the air, and you think you’re absolutely healthy until you begin to sneeze. Your only protection is complete, impossible isolation.”

Jane Jorgensen, principal scientist at Information Extraction & Transport in Arlington, Va., which researches Internet epidemiology for the Defense Advanced Research Projects Agency, isn’t describing the latest flu outbreak but an affliction that affects the Web. One such computer disease emerged this past July and August, and it has computer security researchers more worried about the integrity of the Internet than ever before. The consternation was caused by Code Red, a Web worm, an electronic ailment akin to computerized snakebite. Code Red infects Microsoft Internet Information Servers (IIS). Whereas home computers typically use other systems, many of the most popular Web sites run on IIS. In two lightning-fast strikes, Code Red managed to infiltrate hundreds of thousands of IIS servers in only a few hours, slowing the Internet’s operations. Although Code Red’s effects have waned, patching the security holes in the estimated six million Microsoft IIS Web servers worldwide and repairing the damage inflicted by the worm have cost billions of dollars.

What really disturbs system administrators and other experts, however, is the idea that Code Red may be a harbinger of more virulent Internet plagues. In the past, Web defacements were perpetrated by people breaking into sites individually—the cyberwarfare equivalent of dropping propaganda leaflets on targets. But computer researchers dread the arrival of better-designed automated attack worms that could degrade or even demolish the World Wide Web.

Further, some researchers worry that Code Red was merely a test of the type of computer programs that any government could use to crash the Internet in times of war. This past spring’s online skirmishes over the U.S. spy plane incident with China emphasize the dangers. Full-scale cyberwarfare could cause untold damage to the industrialized world [see “What Happens if the Internet Crashes?” on page 27]. These secret assaults could even enlist your PC as a pawn, making it a “zombie” that participates in the next round of computerized carnage.

BY CAROLYN MEINEL • Photographs by Ethan Hill



Could the Internet crash? This summer’s Code Red attacks could foreshadow destructive cyberwarfare between hacker groups or between governments



AMERICAN HACKERS are being enlisted to help fight the U.S. government's cyberwars.

Save for the scales on which these computer assaults are waged, individual hacking and governmental cyberwarfare are essentially two sides of the same electronically disruptive coin. Unfortunately, it's hard to tell the difference between them until it's too late.

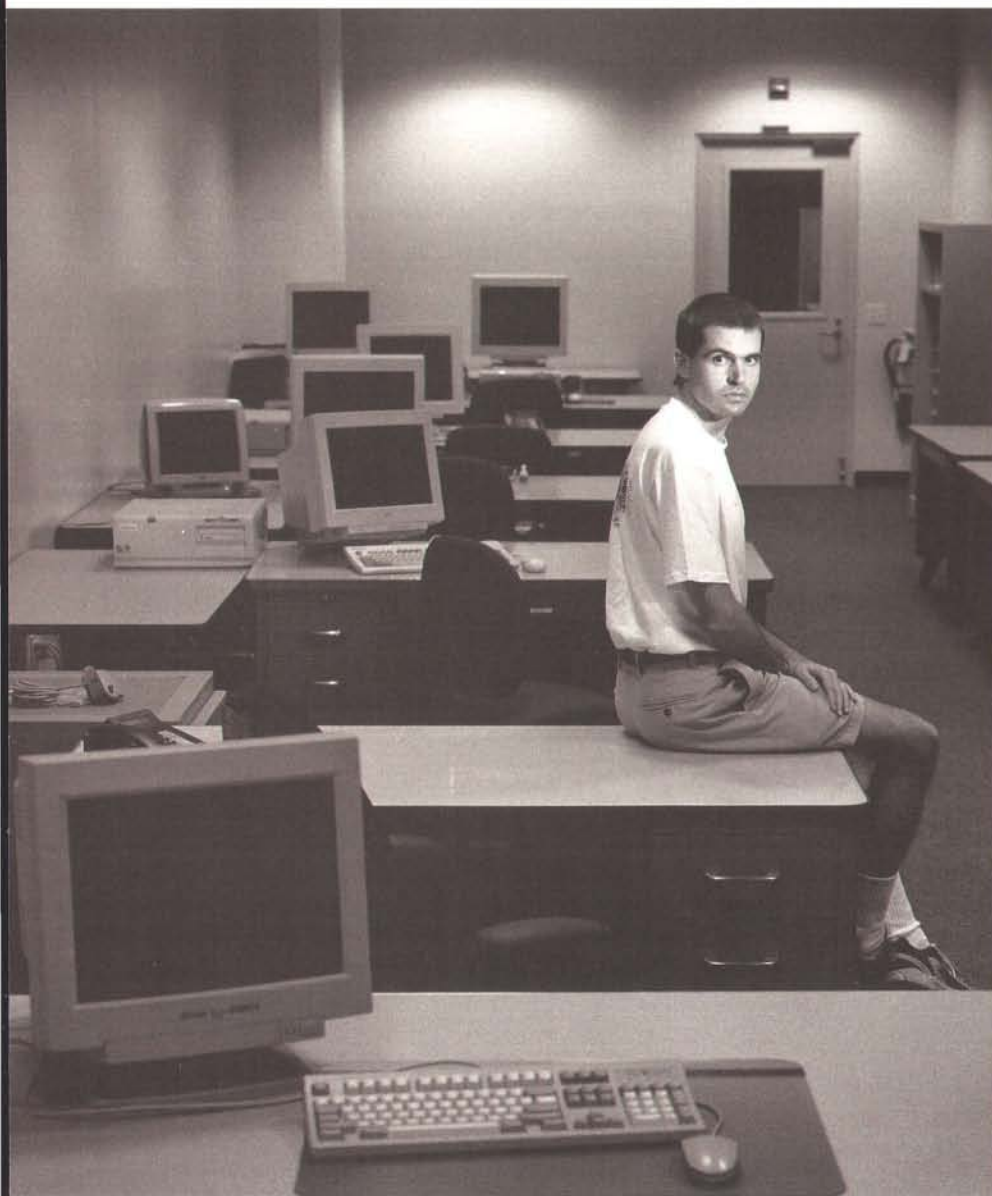
Often popularly lumped in with viruses, Code Red and some similar pests such as Melissa and SirCam are more accurately called worms in the hacker lexicon. Mimicking the actions of its biolog-

ical namesake, a software virus must incorporate itself into another program to run and replicate. A computer worm differs in that it is a self-replicating, self-contained program. Worms frequently are far more infectious than viruses. The Code Red worm is especially dangerous because it conducted what are called distributed denial of service (DDoS) attacks, which overwhelm Internet computers with a deluge of junk communications.

During its July peak, Code Red men-

aced the Web by consuming its bandwidth, or data-transmission capacity. "In cyberwarfare, bandwidth is a weapon," says Gregory Peck, a senior security engineer for FC Business Systems in Springfield, Va., which works to defend U.S. government clients against computer crime. In a DDoS attack, a control computer commands many zombies to throw garbage traffic at a victim in an attempt to use up all available bandwidth. This kind of assault first made the news last year

More than 359,000 servers were infected with the **CODE RED WORM** in less than **14 HOURS**.



when DDoS attacks laid low Yahoo, eBay and other dot-coms.

These earlier DDoS incidents mustered just hundreds to, at most, thousands of zombies. That's because attackers had to break into each prospective zombie by hand. Code Red, being a worm, spreads automatically—and exponentially. This feature provides it with hundreds of times more zombies and hence hundreds of times more power to saturate all available Internet bandwidth rapidly.

The initial outbreak of Code Red contagion was not much more than a case of the sniffles. In the five days after it appeared on July 12, it reached only about 20,000 out of the estimated half a million susceptible IIS computer servers. It wasn't until five days afterward that Ryan Permeh and Marc Maiffret of eEye Digital Security in Aliso Viejo, Calif., a supplier of security software for Microsoft servers, discovered the worm and alerted the world to its existence.

On July 19 the worm reemerged in a more venomous form. "More than 359,000 servers were infected with the Code Red worm in less than 14 hours," says David Moore, senior technical manager at the Cooperative Association for Internet Data Analysis in La Jolla, Calif., a government- and industry-supported organization that surveys and maps the Net's server population. The traffic jam

WEB WATCHER David Moore monitored the rapid spread of Code Red.

What Happens if the Internet Crashes?

WHAT WOULD BE the consequences if the Internet failed in the face of a hacking onslaught? They would be far worse than not being able to make bids on eBay—potentially affecting product manufacturing and deliveries, bank transactions, telephony and more. Should it occur five years from now, the results could be a lot more severe.

Today many businesses use the World Wide Web to order parts and arrange shipments. A collapse of the system would interrupt just-in-time manufacturing, in which components reach the production line within a day or two of being used, to save on inventory costs. Many retail stores also rely on the Web to keep their shelves stocked. Within days, they could start to empty.

By then you may not be able to use your checkbook or ATM card either, as many banks are using the Internet instead of dedicated lines to save money. Other economic institutions such as Wall Street are said to be more susceptible to hackers corrupting trading data than to a shutdown of the system.

The latter eventuality would be met by closing down the market.

Whereas most phones would still work if the Web went down today, experts say that may change a few years from now. Internet telephony started as a way for geek hobbyists to get free long-distance phone calls. Now, however, many calls that originate from an ordinary phone travel part of the way over the public Internet.

Meanwhile unclassified communications of the U.S. Armed Services go through NIPRNET (Non-Secure Internet Protocol Router Network), which uses public Internet communications. The Department of Defense is now "immensely dependent" on NIPRNET, according to Gregory Peck, a senior security engineer for FC Business Systems in Springfield, Va., which provides computer services to the federal government.

Many people ask whether airliners might start falling out of the sky if the World Wide Web crashes. The Federal Aviation Administration's air-traffic control system is sufficiently antiquated that it is in no danger of being held hostage to the Internet.—C.M.

generated by so many computers attempting to co-opt other machines began to overload the capacity of the Internet. By midafternoon, the Internet Storm Center at incidents.org—the computer security industry's watchdog for Internet health—was reporting "orange alert" status. This is one step below its most dire condition, red alert, which signals a breakdown.

Then, at midnight, all Code Red zombies quit searching for new victims. Instead they all focused on flooding one of the servers that hosts the White House Web site with junk connections, threatening its shutdown. "The White House essentially turned off one of its two DNS servers, saying that any requests to whitehouse.gov should be rerouted to the other server," says Jimmy Kuo, a Network Associates McAfee fellow who assisted the White House in finding a solution. Basically, the system administrators dumped all communications addressed to the compromised server. As it turned out, Code Red couldn't cope with the altered Internet protocol address and waged war on the inactive site. "The public didn't notice anything, because any requests went to the other server," Kuo says.

By the close of July 20, all existing Code Red zombies went into a preprogrammed eternal sleep. As the worms lodge only in each computer's RAM memory, which is purged when the ma-

chine shuts down, all it took was a reboot to eradicate their remnants. Case closed.

Or was it? A few days later analysts at eEye revealed that if someone were to release a new copy of Code Red at any time between the first through the 19th day of any month (the trigger dates coded in by the original hacker), the infection would take off again.

Over the next 10 days computer security volunteers worked to notify Microsoft IIS users of the vulnerability of their servers. On July 29 the White House held a press conference to implore people to protect their IIS servers against Code Red's attacks. "The mass traffic associated with this worm's propagation could degrade the functioning of the Internet," warned Ronald L. Dick, director of the FBI's National Infrastructure Protection Center. By the next day Code Red was all over the news.

The second coming of Code Red was, as expected, weaker than the first. On August 1, it infected approximately 175,000 servers—nearly all those susceptible and about half the total of the previous episode. A slower infection rate and fewer vulnerable servers held Internet disruptions to a minimum. After a while, the second attack subsided.

But that was not the end. Yet another worm was unleashed on August 4 using the same break-in method as Code Red. The new worm, dubbed Code Red

II, installed a backdoor allowing a master hacker to direct the activities of victim computers at will. The worm degraded intranets with "arp storms" (floods of Ethernet packets) and hunted for new victims. In short order, Code Red II disabled parts of the Web-based e-mail provider Hotmail, several cable and digital subscriber-line (DSL) Internet providers and part of the Associated Press news distribution system. As time passed, Code Red II managed to infect many corporate and college intranets. Halfway through August, Code Red II disabled some Hong Kong government internal servers. The most common victim computers were personal Web servers run by Windows 2000 Professional. This rash of disruptions prompted incidents.org to again declare an orange alert. Experts estimate that 500,000 internal servers were compromised.

In mid-August, Computer Economics, a security research company, said that Code Red had cost \$2 billion in damage. By the time it is fully purged from the Internet, the computer attack will probably rank among the most expensive in histo-

THE AUTHOR

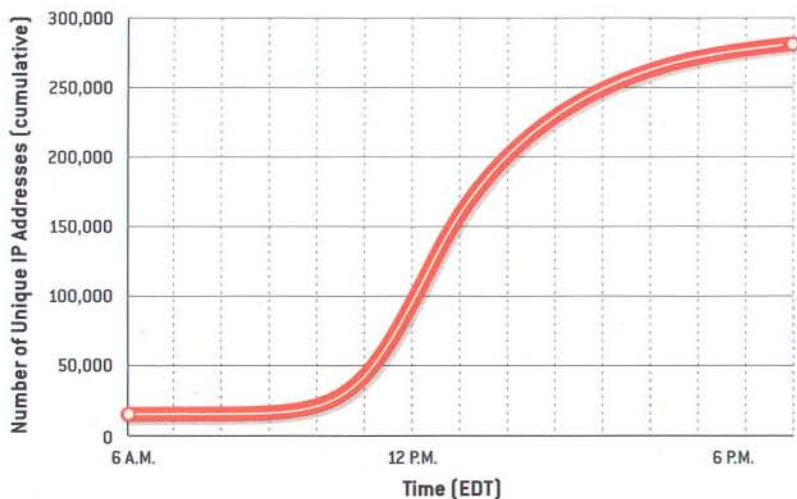
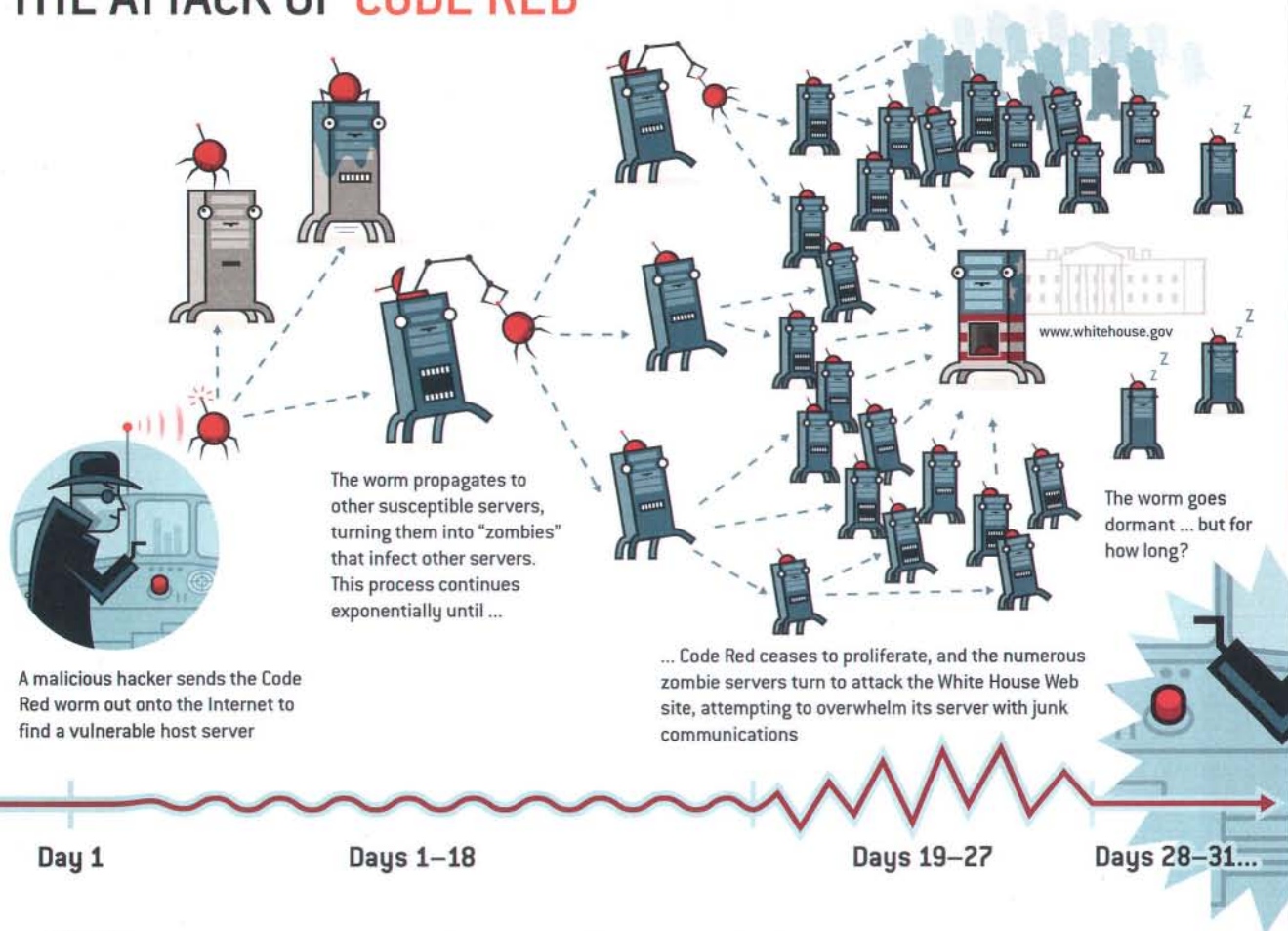
CAROLYN MEINEL writes frequently about computer security. Based in Sandia Park, N.M., she is the author of *The Happy Hacker and Überhacker! How to Break into Computers*. Meinel's upcoming book, *War in Cyberspace*, examines Internet warfare. Her Web site, happyhacker.org, is a resource for home computer users.

Code Red

internet worms

CODE RED is an Internet worm that infects unprotected Microsoft Internet Information Servers (IIS), on which many popular Web sites run. During the summer, the worm's secret assaults turned IIS computers into "zombies" that conducted what is called a distributed denial of service attack on the White House Web site, attempting to overwhelm it by flooding it with garbage communications. More effective worms have the potential to saturate the Web's data-transmission capacity, possibly disabling the Internet.

THE ATTACK OF CODE RED



INTERNET PROTOCOL ADDRESSES INFECTED BY CODE RED

RAPID RISE—During a 12-hour period on July 19, 2001, the number of Internet protocol addresses compromised by the first large-scale assault of the Code Red worm surged from around 16,000 to about 280,000. After its initial spread, Code Red went dormant. Soon thereafter, however, a reinfection caused another, smaller outbreak. Experts estimate that the worm's attacks and the following Code Red II outburst will cost several billion dollars to rectify.

ry. Nearly \$9 billion was spent to fight last year's LoveLetter virus, and 1999's Melissa worm assault cost \$1 billion to repair.

Of course, Code Red isn't the only worm out there. Some of them are aimed at home computers. A worm called W32/Leaves, for example, permits a remote attacker to control infected PCs in a coordinated fashion, enabling synchronized waves of attacks. (Although Code Red II allows this possibility as well, it lacks the coding that enables remote control.) The Computer Emergency Response Team, a federally funded watchdog organization at Carnegie Mellon University, has received reports of more than 23,000 W32/Leaves zombies. The current total is unknown, but as W32/Leaves continues to propagate, the infected population will probably grow significantly. In July, Britain's Scotland Yard charged an unidentified 24-year-old man with creating W32/Leaves.

"Almost any computer, operating

a U.S. Navy EP-3E spy plane this past April give a hint of how such a conflict might play out.

According to accounts in the press, the hacker exchanges began when negotiations for the release of American hostages stalled. On April 9 and 10, attackers defaced two Chinese Web sites with slurs, insults and even threats of nuclear war. During the following week, American hackers hit dozens more Chinese sites. Those supporting China responded by disfiguring one obscure U.S. Navy Web site.

China, however, held a weapon in reserve. In late March the National Infrastructure Protection Center had warned of a new worm on the loose: the Ii0n Worm. Lion, the hacker who founded the hacker group H.U.C. (Honkers Union of China), has taken credit for writing it. Unlike the initial Code Red's preprogrammed zombies, Ii0n's zombies accept new commands from a central computer. Also, Ii0n infects Linux computers, which

also urged to call off all irrational actions and turn their enthusiasm into strength to build up the country and safeguard world peace."

U.S. law-enforcement agencies, the White House and U.S. hacker organizations never objected to the American side of this cyberconflict, although the FBI's infrastructure center had warned of "the potential for increased hacker activity directed at U.S. systems."

How to Wage Covert Cyberwarfare

IN VIEW OF the spy-plane episode, some commentators have wondered whether the U.S. federal government encouraged American hackers to become agents of cyberwar. After all, the U.S. has worked with private groups to wage covert warfare before, as in the Iran/Contra scandal. And links between the two communities have been reported. It's difficult, however, to say exactly how strong the connection between hackers and the government

Code Red II installed a backdoor allowing a

MASTER HACKER to direct victim computers at will.

system or software you may buy contains weaknesses that the manufacturer knows lets hackers break in," says Larry Leibrock, a leading researcher in computer forensics and associate dean for technology of the business school of the University of Texas at Austin. Future "federal regulation could require that vendors take the initiative to contact customers and help them upgrade their products to fix security flaws," he continues. "Today, however, it is up to each consumer to hunt down and fix the many ways hackers and cyberwarriors exploit to abuse their computers."

World Cyberwars

BEYOND THE THREAT posed by malicious hacker programs is the danger of Internet attacks conducted in a concerted fashion by top computer talent spurred to act by international events. The cyberbattles that broke out over the collision of a Chinese fighter plane that collided with

means it can masquerade as any computer on the Net. This property makes it hard to track down infected servers.

Meanwhile pro-U.S. hack attacks escalated. The official Chinese publication, *People's Daily*, reported that "by the end of April over 600 Chinese Web sites had come under fire." In contrast, Chinese hackers had hit only three U.S. sites during the same period.

In the next few days the Chinese hacker groups H.U.C., Redcrack, China Net Force, China Tianyu and Redhackers assaulted a dozen American Web sites with slogans such as "Attack anti-Chinese arrogance!" On the first of May several DDoS strikes were initiated. Over the next week Chinese hackers took credit for wrecking about 1,000 additional American Web sites.

On May 7 China acknowledged its responsibility for the DDoS attacks and called for peace in a *People's Daily* news story. It ran: "The Chinese hackers were

might be. Clearly, the murky world of hacking doesn't often lend itself to certainty. And because it is the policy of the U.S. National Security Agency and various Defense Department cyberwarfare organizations not to comment on Web security matters, these relationships cannot be confirmed. Still, the indications are at least suggestive.

Consider the history of Fred Vilella, now an independent computer consultant. According to numerous press reports and his own statements, Vilella took part in counterterrorism activities in the 1970s. In 1996 he hired hackers of the Dis Org Crew to help him conduct training sessions on the hacker threat for federal agencies. This gang also helps to staff the world's largest annual hacker convention, Def Con.

Erik Ginorio (known to the hacker world as Bronc Buster) publicly took credit for defacing a Chinese government Web site on human rights in October

1998. This act is illegal under U.S. law. Not only was Ginorio not prosecuted, he says Villella offered him a job. Villella could not be reached for comment.

In another hacker-government connection, Secure Computing in San Jose, Calif., became a sponsor of Def Con in 1996. According to its 10-K reports to the U.S. Securities and Exchange Commission, Secure Computing was created at the direction of the National Security Agency, the supersecret code-breaking and surveillance arm of the U.S. government. Two years after that, Secure Computing hired the owner of Def Con, Jeff Moss. Several former Villella instructors also staffed and managed Def Con.

Questionable things happen at Def Cons. At the 1999 Def Con, for example, the Cult of the Dead Cow, a hacker gang

headquartered in Lubbock, Tex., put on a mediagenic show to promote its Back Orifice 2000 break-in program. Gang members extolled the benefits of "hacking to change the world," claiming that eight-year-olds could use this program to break into Windows servers.

Meanwhile Pieter Zatkó, a Boston-area hacker-entrepreneur and a member of the gang, was onstage promoting a software plug-in for sale that increased the power of Back Orifice 2000. According to the Cult's Web site, Back Orifice 2000 was downloaded 128,776 times in the following weeks. On February 15, 2000, President Bill Clinton honored Zatkó for his efforts by inviting him to the White House Meeting on Internet Security. Afterward Zatkó remained with a small group to chat with the president.

Every year Def Con holds a "Meet the Feds" panel. At its 2000 meeting, Arthur L. Money, former U.S. assistant secretary of defense for command, control, communications and intelligence, told the crowd, "If you are extremely talented and you are wondering what you'd like to do with the rest of your life—join us and help us educate our people [government personnel]."

In 1997 Moss launched the Black Hat Briefings. In hacker lingo, a black hat is a computer criminal. Theoretically, these meetings are intended to train people in computer security. They bear considerable similarity to Def Con, however, only with a \$1,000 price tag per attendee. Their talks often appear to be more tutorials in how to commit crime than defend against it. For example, at one session attendees

What Can Be Done to Defend the Web?

AS POGO the comic-strip character said, "We have met the enemy and he is us." One of the weakest links in protecting the Internet is the home PC user. Cybernetic worms—self-replicating hacker software that can wreak havoc on Internet operations—can turn personal computers into "zombies," or slave agents that help to destroy other computer operations. Of particular concern are worms that can conduct effectively targeted distributed denial of service (DDoS) attacks, in which zombie computers deluge a Web site with useless communications.

Computer professionals are being asked to get the word out to home users to check for zombies. "That's because our worst Internet nightmare is the grandma who uses her DSL [high-bandwidth-capacity digital subscriber line that is always connected] to shop on eBay," says Gregory Peck of FC Business Systems, which provides computer services to the federal government. High bandwidth means that a home zombie can pump lots of junk into the Internet, swamping targeted Web sites.

You may think your home computer is safe from assault because it runs automatic virus updates or because you registered your software and receive vendor e-mails about product upgrades. Guess again. Few vendors feel obligated to help users keep hackers out. That's why it's important for home users to install firewalls.

Complicating the safety issue, most new PCs will soon be running the Windows XP operating system, which enables "raw sockets." Sockets are software constructions that generate the packets [the smallest data-transmission units] that transfer information across networks. With raw sockets technology, packets can be crafted in an arbitrary manner even if that violates safeguarding protocols. Raw sockets, for example,

enabled the 110n worm to hide on Linux servers by forging Internet addresses [see preceding page]. They also allow hackers to create malformed packets that will crash a receiving computer.

Beyond the home PC, another approach to defending the Web is to arrest more computer criminals. Nowadays, though, dangerous attackers may operate through a chain of compromised computers, with one or more being located across national borders. To obtain evidence in these cases requires cooperation among the law-enforcement agencies of two or more countries.

International pursuit of computer criminals would be made easier by adoption of the "Convention on Cyber-crime" now under consideration by the 44 nations of the Council of Europe, which includes the U.S., Canada and Japan. Part of the treaty would also criminalize possession or creation of computer crime instructions or programs except for the authorized testing or protection of a computer system. [The text of the Cyber-crime Treaty is available at conventions.coe.int/treaty/EN/projets/cybercrime.htm]

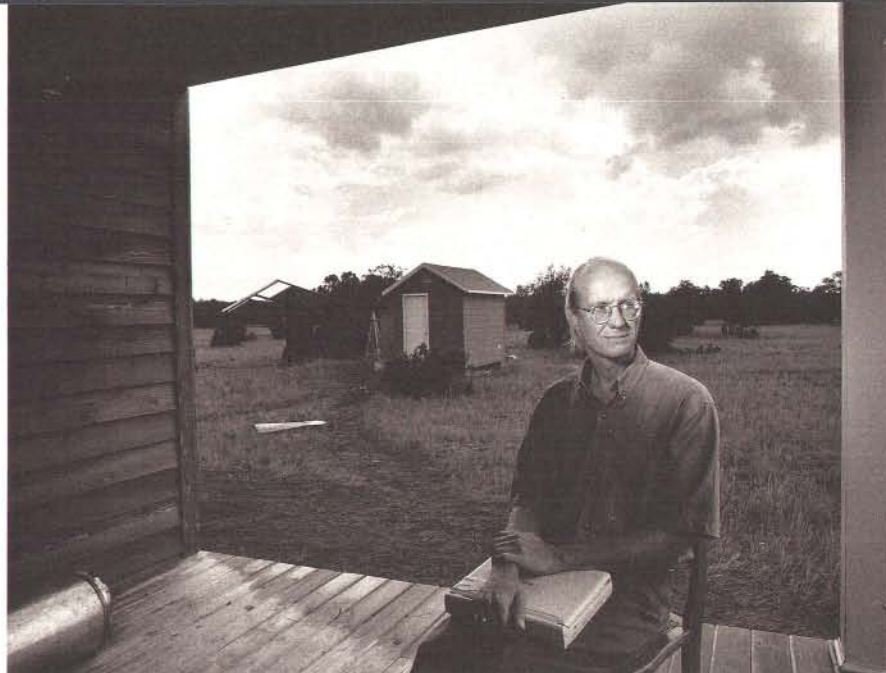
These restraints are controversial, though. At least 35 lobby groups, including the Electronic Frontier Foundation and the Global Internet Liberty Campaign, oppose the treaty because they believe it would restrict freedom of speech and invade personal privacy. It's hard to find antidotes to viruses and worms if researchers cannot study copies of them on their computers.

Another solution is to require that Internet servers be secure. For example, the U.S. Federal Trade Commission proposed a regulation in July that requires financial service companies to guard their networks against "anticipated threats." This is only a small step in the right direction. —C.M.

NET VIROLOGIST Mark A. Ludwig
writes about computer viruses and worms
at his rural Arizona home.

learned about "Evidence-Eliminator," billed as being able to "defeat the exact same forensic software as used by the U.S. Secret Service, Customs Department and Los Angeles Police Department."

It should be noted that the U.S. government does have a formal means to wage cyberwar. On October 1, 2000, the U.S. Space Command took charge of the Computer Network Attack mission for the Department of Defense. In addition, the U.S. Air Force runs its Information Warfare Center research group, located in San Antonio.



Get enough zombies attacking enough targets, and the **ENTIRE INTERNET** could become unusable.

Given these resources, why would the U.S. and China encourage cybermilitias? "It's very simple. If you have an unofficial army, you can disclaim them at any time," says Mark A. Ludwig, author of *The Little Black Book of Computer Viruses* and the upcoming *The Little Black Book of Internet Viruses*. "If your military guys are doing it and you are traced back, the egg's on your face."

Wherever it came from, the Code Red assault was just a taste of what a concerted cyberwar could become. "I think we can agree that it was not an attempt at cyberwar. The worm was far too noisy and easily detected to be much more than graffiti/vandalism and a proof-of-concept," says Harlan Carvey, an independent computer security consultant based in Virginia.

Stuart Staniford, president of Silicon Defense in Eureka, Calif., notes, however, that if the zombie computers "had a long target list and a control mechanism to allow dynamic retargeting, [they] could have DDOSed [servers] used to map addresses to contact information, the ones used to distribute patches, the ones belonging to companies that analyze worms or distribute incident response information. Code Red illustrates that it's not much harder for a worm to get *all* the vul-

nerable systems than it is to get some of them. It just has to spread fast enough."

Code Red already offers deadly leverage for nefarious operators, according to Marc Maiffret, who bills himself as "chief hacking officer" of eEye: "The way the [Code Red] worm is written, it could allow online vandals to build a list of infected systems and later take control of them."

Get enough zombies attacking enough targets, and the entire Internet could become unusable. Even the normal mechanisms for repairing it—downloads of instructions and programs to fix zombies and the ability to shut off rogue network elements—could become unworkable. In addition, hackers constantly publicize new ways to break into computers that could be used by new worms. A determined attacker could throw one devas-

tating worm after another into the Internet, hitting the system every time it struggled back and eventually overpowering it.

"We know how [crashing the Internet] can be done right," says Richard E. Smith, a researcher with Secure Computing and author of the newly published book *Authentication*. "What I've found particularly disquieting is how little public fuss there's been [about Code Red]. The general press has spun the story as being an unsuccessful attack on the White House as opposed to being a successful attack on several hundred thousand servers: 'Ha, ha, we dodged the bullet!' A cynic might say this demonstrates how 'intrusion tolerant' IIS is—the sites are all penetrated but aren't disrupted enough to upset the owners or generate much press comment. The rest of us are waiting for the other shoe to drop."

MORE TO EXPLORE

The Computer Emergency Response Team's Guide to Home Network Security:
www.cert.org/tech_tips/home_networks.htm

The Internet Storm Center: www.incidents.org

The National Infrastructure Protection Center: www.nipcc.gov

The Cooperative Association for Internet Data Analysis: www.caidda.org

Microsoft Windows NT, 2000 and XP security information: www.ntbugtraq.org

Free security test for home computers: grc.com and security2.norton.com/us/home.asp

Microsoft Windows NT, 2000 and XP information:

www.microsoft.com/technet/treeview/default.asp?url=/technet/itsolutions/security/current.asp

SCIENTIFIC AMERICAN

Help Desk

For your convenience, we have summarized the answers to often-asked customer service questions and how to get in touch with us.

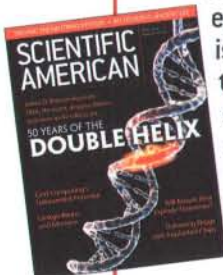
www.sciam.com

Our award-winning Internet resource, updated daily Monday-Friday and thoroughly linked to related sites. And remember, www.sciam.com can help you:

- Enter changes of address
- Renew subscriptions
- Enter new orders
- Answer your questions

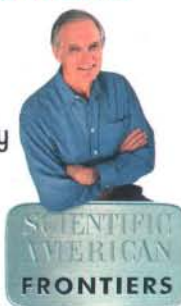
Scientific American

SCIENTIFIC AMERICAN is considered the authoritative source, keeping its readers up-to-date while enriching their world. Your 12-issue subscription includes the annual September single-topic issue. Annual subscriptions: In the U.S., Scientific American is \$34.97. In Canada, \$49. Elsewhere, \$55.



Scientific American Frontiers

The magazine's PBS television series, Scientific American Frontiers, is hosted by actor and lifelong science buff Alan Alda. There are six one-hour episodes each season. Visit the Scientific American Frontiers Web site at www.pbs.org/saf



Subscription inquiries

www.sciam.com is our secured website where you can subscribe online or call us: In the U.S. and Canada, 1-800-333-1199. Elsewhere 1-515-247-7631. Email: subscriptions@sciam.com, or write: Subscription Manager, Scientific American, PO Box 3187, Harlan, IA 51537. You can locate the date of the last issue of your subscription on the mailing label. For changes of address, please notify us at least eight weeks in advance, and include both old and new addresses.

Scientific American Digital

www.sciamdigital.com is more than just a digital magazine. You also have access to over 120 issues of SCIENTIFIC AMERICAN from 1993 to the present. SCIENTIFIC AMERICAN DIGITAL is \$34.95 per year and includes all Exclusive Online Issues for free. Special Editions are available for purchase at \$5 each.

Back issues

In U.S., \$10.95 each. Elsewhere, \$13.95 each. Many issues are available. Fax your order with your Visa, MasterCard or AMEX information to (212) 355-0408.

Reprints

\$4 each, minimum order 25 copies, prepaid. Limited availability. Write to Reprint Department, Scientific American, 415 Madison Avenue, New York, NY 10017-1111; fax: 212-355-0408 or tel: 212-451-8877. Email: reprints@sciam.com

Photocopying rights

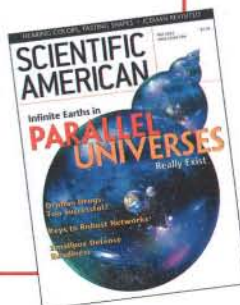
Granted by Scientific American, Inc., to libraries and others registered with the Copyright Clearance Center (CCC) to photocopy articles in this issue for the fee of \$3.50 per copy of each article plus \$0.50 per page. Such clearance does not extend to photocopying articles for promotion or other commercial purpose.

Subscriber alert!

SCIENTIFIC AMERICAN has been made aware that some subscribers have received notifications/subscription offers from companies such as Publishers Services Exchange, United Publishers Network, Global Publication Services, Publishers Access Service and Lake Shore Publishers Service. These are not authorized representatives/agents of Scientific American and are selling subscriptions at a much higher rate than the regular subscription or renewal price. Please forward any correspondence you may receive from these companies to:

L. Terlecki
Scientific American
415 Madison Ave.
New York, NY 10017

SCIENTIFIC AMERICAN



The Incredible Shrinking Circuit

RESEARCHERS HAVE BUILT
NANOTRANSISTORS AND NANOWIRES.
NOW THEY JUST NEED TO FIND A WAY
TO PUT THEM ALL TOGETHER

BY CHARLES M. LIEBER

NANOWIRES, each about five to 10 nanometers in diameter, may represent the future of electronics. They are the brown lines, made of indium phosphide, connecting the gold electrodes in this micrograph. These wires have been put to truly diverse uses—as memory and logic and as arrays of light-emitting diodes.

Do we really need to keep on making circuits smaller? The miniaturization of silicon microelectronics seems so inexorable that the question seldom comes up—except maybe when we buy a new computer, only to find that it becomes obsolete by the time we leave the store. A state-of-the-art microprocessor today has more than 40 million transistors; by 2015 it could have nearly five billion. Yet within the next two decades this dramatic march forward will run up against scientific, technical and economic limits. A first reaction might be, So what? Aren't five billion transistors enough already?

Yet when actually confronted with those limits, people will no doubt want to go beyond them. Those of us who work to keep computer power growing are motivated in part by the sheer challenge of discovering and conquering unknown territory. But we also see the potential for a revolution in medi-

els of organization. The basic building block is usually the transistor or its nanoequivalent—a switch that can turn an electric current on or off as well as amplify signals. In microelectronics, transistors are made out of chunks of semiconductor—a material, such as impure silicon, that can be manipulated to flip between conducting and nonconducting states. In nanoelectronics, transistors might be organic molecules or nanoscale inorganic structures.

The next level of organization is the interconnection—the wires that link transistors together in order to perform arithmetic or logical operations. In microelectronics, wires are metal lines typically hundreds of nanometers to tens of microns in width deposited onto the silicon; in nanoelectronics, they are nanotubes or other wires as narrow as one nanometer.

At the top level is what engineers call architecture—the over-

The use of molecules for electronic devices remained theoretical until a recent confluence of advances in chemistry, physics and engineering.

cine and so many other fields, as extreme miniaturization enables people and machines to interact in ways that are not possible with existing technology.

As the word suggests, microelectronics involves components that measure roughly one micron on a side (although lately the components have shrunk to a size of almost 100 nanometers). Going beyond microelectronics means more than simply shrinking components by a factor of 10 to 1,000. It also involves a paradigm shift for how we think about putting everything together.

Microelectronics and nanoelectronics both entail three lev-

els of organization. The basic building block is usually the transistor or its nanoequivalent—a switch that can turn an electric current on or off as well as amplify signals. In microelectronics, transistors are made out of chunks of semiconductor—a material, such as impure silicon, that can be manipulated to flip between conducting and nonconducting states. In nanoelectronics, transistors might be organic molecules or nanoscale inorganic structures.

The next level of organization is the interconnection—the wires that link transistors together in order to perform arithmetic or logical operations. In microelectronics, wires are metal lines typically hundreds of nanometers to tens of microns in width deposited onto the silicon; in nanoelectronics, they are nanotubes or other wires as narrow as one nanometer. At the top level is what engineers call architecture—the over-

Overview/*Nanoelectronics*

- Silicon chips, circuit boards, soldering irons: these are the icons of modern electronics. But the electronics of the future may look more like a chemistry set. Conventional techniques can shrink circuits only so far; engineers will soon need to shift to a whole new way of organizing and assembling electronics. One day your computer may be built in a beaker.
- Researchers have created nanometer-scale electronic components—transistors, diodes, relays, logic gates—from organic molecules, carbon nanotubes and semiconductor nanowires. Now the challenge is to wire these tiny components together.
- Unlike conventional circuit design, which proceeds from blueprint to photographic pattern to chip, nanocircuit design will probably begin with the chip—a haphazard jumble of as many as 10^{24} components and wires, not all of which will even work—and gradually sculpt it into a useful device.

Smallifying Machines

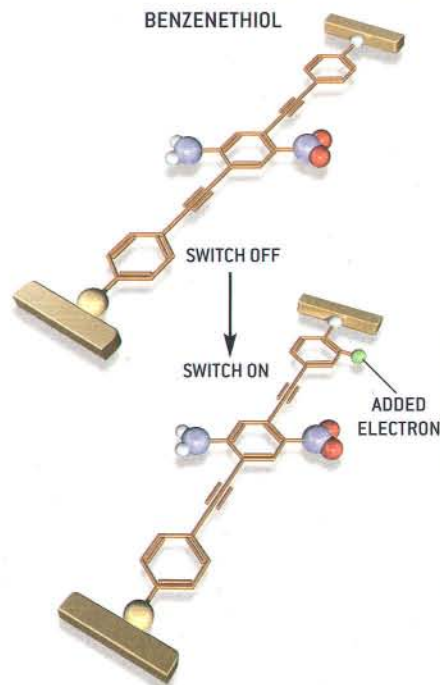
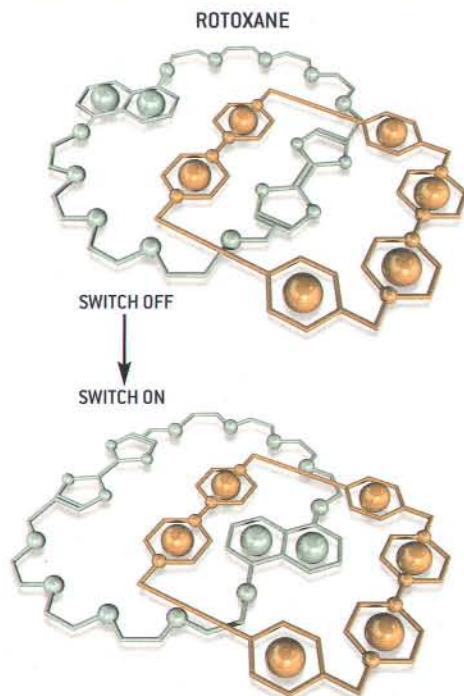
THE USE OF MOLECULES for electronic devices was suggested more than a quarter of a century ago in a seminal paper by Avi Aviram of IBM and Mark A. Ratner of Northwestern University. By tailoring the atomic structures of organic molecules, they proposed, it should be possible to concoct a transistorlike device. But their ideas remained largely theoretical until a recent confluence of advances in chemistry, physics and engineering.

Of all the groups that have turned Aviram and Ratner's

NANOTRANSISTORS

MOLECULAR TRANSISTORS

could be the building blocks of electronics on the nanometer scale. Each of the two molecules shown here conducts electricity like a tiny wire once a chemical reaction—oxidation reduction—alters its atomic configuration and switches it on. In the diagram, each stick represents a chemical bond; each intersection of two sticks represents a carbon atom; and each ball represents an atom other than carbon.



ideas into reality, two teams—one at the University of California at Los Angeles and Hewlett-Packard, the other at Yale, Rice and Pennsylvania State universities—stand out. Within the past year, both have demonstrated that thousands of molecules clustered together can carry electrons from one metal electrode to another. Each molecule is about 0.5 nanometer wide and one or more nanometers long. Both groups have shown that the clusters can behave as on/off switches and might thus be usable in computer memory; once on, they will stay on for 10 minutes or so [see “Computing with Molecules,” by Mark A. Reed and James M. Tour; *SCIENTIFIC AMERICAN*, June 2000]. That may not sound like a long time, but computer memory typically loses its information instantly when the power is turned off; even when the power is on, the stored information leaks away and must be “refreshed” every 0.1 second or so.

Although the details differ, the switching mechanism for both molecules is believed to involve a well-understood chemical reaction, oxidation reduction, in which electrons shuffle among atoms within the molecule. The reaction puts a twist in the molecule, blocking electrons as surely as a kink in a hose blocks water [see illustration above]. In the “on” position, the clusters of molecules may conduct electricity as much as 1,000 times better than in the “off” position. That ratio is actually rather low compared with that of typical semiconductor transistors, whose conductivity varies a millionfold. Researchers are now looking for other molecules with even better switching properties and are also working to understand the switching process itself.

My own research group at Harvard University is one of several that have focused not on organic molecules but on long, thin inorganic wires. The best-known example is the carbon nanotube, which is typically about 1.4 nanometers in diameter [see “Nanotubes for Electronics,” by Philip G. Collins and Phaedon Avouris; *SCIENTIFIC AMERICAN*, December 2000]. Not only can these nanoscale wires carry much more current, atom for atom, than ordinary metal wires, they can also act as tiny transistors. By functioning both as interconnections and as components, nanowires kill two birds with one stone. Another advantage is that they can exploit the same basic physics as standard silicon microelectronics, which makes them easier to understand and manipulate.

In 1997 Cees Dekker’s group at the Delft University of Technology in the Netherlands and Paul L. McEuen’s group, then at the University of California at Berkeley, independently reported highly sensitive transistors made from metallic carbon nanotubes. These devices could be turned on and off by a

THE AUTHOR

CHARLES M. LIEBER spent much of his childhood building—and breaking—stereos, cars and model airplanes. He is now the Mark Hyman Professor of Chemistry at Harvard University, where he directs a group of 25 undergraduate, graduate and postdoctoral researchers who focus on nanoscale science and technology. Lieber recently founded NanoSys, Inc., with Larry Bock of CW Ventures and Hongkun Park of Harvard. Their modest goal: to revolutionize chemical and biological sensing, electronics, optoelectronics, and information storage.

single electron but required very low temperatures to operate. This past July Dekker's team swept away this limitation. The researchers used an atomic force microscope to create a single-electron transistor that could function at room temperature. Dekker and his co-workers have also fashioned a more conventional field-effect transistor, the building block of most integrated circuits today, out of a carbon nanotube, and McEuen's group has combined metallic and semiconductor nanotubes into

a diode, which allows electric current to pass in one direction only. Finally, my group has demonstrated a very different type of switch, a nanoscale electromechanical relay.

Hot Wire

A MAJOR PROBLEM with nanotubes is that they are difficult to make uniform. Because a slight variation in diameter can spell the difference between a conductor and a semiconductor,

DNA Computing

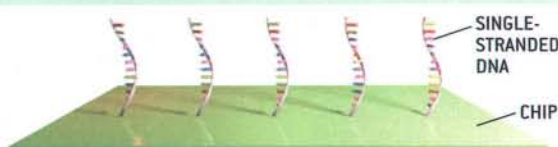
WHY LIMIT OURSELVES TO ELECTRONICS? Most efforts to shrink computers assume that these machines will continue to operate much as they do today, using electrons to carry information and transistors to process it. Yet a nanoscale computer could operate by completely different means. One of the most exciting possibilities is to exploit the carrier of genetic information in living organisms, DNA.

The molecule of life can store vast quantities of data in its sequence of four bases (adenine, thymine, guanine and cytosine), and natural enzymes can manipulate this information in a highly parallel manner. The power of this approach was first brought to light by computer scientist Leonard M. Adleman in 1994. He showed that a DNA-based computer could solve a type of problem that is particularly difficult for ordinary computers—the Hamiltonian path problem, which is related to the infamous traveling-salesman problem [see “Computing with DNA,” by Leonard M. Adleman; *SCIENTIFIC AMERICAN*, August 1998].

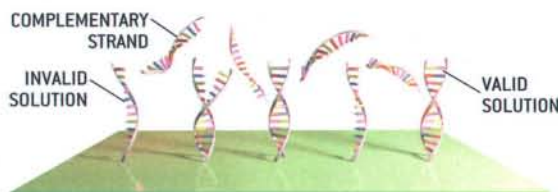
Adleman started by creating a chemical solution of DNA. The individual DNA molecules encoded every possible pathway between two points. By going through a series of separation and amplification steps, Adleman weeded out the wrong paths—those, for example, that contained points they were not supposed to contain—until he had isolated the right one. More recently, Lloyd M. Smith's group at the University of Wisconsin—Madison implemented a similar algorithm using gene chips, which may lend themselves better to practical computing (*diagram*).

Despite the advantages of DNA computing for otherwise intractable problems, many challenges remain, including the high incidence of errors caused by base-pair mismatches and the huge number of DNA nanoelements needed for even a modest computation. DNA computing may ultimately merge with other types of nanoelectronics, taking advantage of the integration and sensing made possible by nanowires and nanotubes.

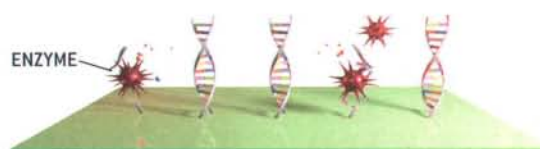
—C.M.L.



- 1 Single DNA strands are attached to a silicon chip. They encode all possible values of the variables in an equation that the researchers want to solve.



- 2 Copies of a complementary strand—which encodes the first clause of the equation—are poured onto the chip. These copies attach themselves to any strand that represents a valid solution of the clause. Any invalid solutions remain a single strand.



- 3 An enzyme removes all the single strands.



- 4 Other processes melt away the added complementary strands. These steps are repeated with all the clauses of the equation.



- 5 The DNA strand that survives this whole process represents the solution to the whole equation.

Soon nanodevices may have useful applications—for example, as ultrasensitive detectors of gas molecules and biological compounds.

a large batch of nanotubes may contain only a few working devices. In April of this year Phaeton Avouris and his colleagues at the IBM Thomas J. Watson Research Center came up with a solution. They started with a mixture of conducting and semiconducting nanotubes and, by applying a current between metal electrodes, selectively burned away the conducting ones until just semiconducting ones were left. The solution is only partial, however, because it requires the use of conventional lithography to wire up the random nanotube array and then test and modify each of the individual elements, which would ultimately number in the billions.

My group has also been working on a different type of nanoscale wire, which we term the semiconductor nanowire. It is about the same size as a carbon nanotube, but its composition is easier to control precisely. To synthesize these wires, we start with a metal catalyst, which defines the diameter of the growing wire and serves as the site where molecules of the desired material tend to collect. As the nanowires grow, we incorporate chemical dopants (impurities that add or remove electrons), thereby controlling whether the nanowires are *n*-type (having extra electrons) or *p*-type (having a shortage of electrons or, equivalently, a surfeit of positively charged “holes”).

The availability of *n*- and *p*-type materials, which are the essential ingredients of transistors, diodes and other electronic devices, has opened up a new world for us. We have assembled a wide range of devices, including both major types of transistors (field-effect and bipolar); inverters, which transform a “0” signal to a “1”; and light-emitting diodes, which pave the way for optical interconnections. Our bipolar transistors were the first molecular-scale devices ever to amplify a current. A recent advance in my lab by Xiangfeng Duan has been the assembly of memory from crisscrossing *n*- and *p*-type nanowires. The memory can store information for 10 minutes or longer by trapping charge at the interface between the crossing nanowires [see illustration on next page].

Breaking the Logjam

BUILDING UP AN ARSENAL of molecular and nanoscale devices is just the first step. Interconnecting and integrating these devices is perhaps the much greater challenge. First, the nanodevices must be connected to molecular-scale wires. To date, organic-molecule devices have been hooked up to conventional metal wires created by lithography. It will not be easy to substitute nanowires, because we do not know how to make a good electrical connection without ruining these tiny wires in the process. Using nanowires and nanotubes both for the devices and for the interconnections would solve that problem.

Second, once the components are attached to nanowires, the

wires themselves must be organized into, for example, a two-dimensional array. In a report published earlier this year, Duan and another member of my team, Yu Huang, made a very significant breakthrough: they assembled nanocircuits by means of fluid flows. Just as sticks and logs can flow down a river, nanoscale wires can be drawn into parallel lines using fluids. In my lab we have used ethanol and other solutions and controlled the liquid flow by passing it through channels molded into polymer blocks, which can be easily placed on the substrate where we wish to assemble devices.

The process creates interconnections in the direction of the fluid flow: if the flow is along only one channel, then parallel nanowires are formed. To add wires in other directions, we redirect the flow and repeat the process, building up additional layers of nanowires. For instance, to produce a right-angle grid, we first lay down a series of parallel nanowires, then rotate the direction of flow by 90 degrees and lay down another series. By using wires of different compositions for each layer, we can rapidly assemble an array of functional nanodevices using equipment not much more sophisticated than a high school chemistry lab. A grid of diodes, for example, consists of a layer of conducting nanotubes above a layer of semiconductor nanotubes, or a layer of *n*-type nanowires atop a layer of *p*-type nanowires. In both cases, each junction serves as a diode.

Our approach, which is similar to that being pursued by the team at U.C.L.A. and Hewlett-Packard, is deterministic. We are trying to create arrays with a certain predictable behavior. Form follows function. An alternative proposed by the group at Rice, Yale and Penn State is to allow blocks of devices and wires to interconnect at random. Later, the ensemble can be analyzed to determine how it might be used for storage or computation. In this case, function follows form. The problem with this procedure is that it would take a huge effort to map a complex network and figure out what use it could be put to.

Intimately linked to all these efforts is the development of architectures that best exploit the unique features of nanoscale devices and the capabilities of bottom-up assembly. Although we can make unfathomable numbers of dirt-cheap nanostructures, the devices are much less reliable than their microelectronic counterparts, and our capacity for assembly and organization is still quite primitive.

In collaboration with André DeHon of the California Institute of Technology, my group has been working on highly simplified architectures that can be generalized for universal computing machines. For memory, the architecture starts with a two-dimensional array of crossed nanowires or suspended electromechanical switches in which one can store information at each cross point. The same basic architecture is being pur-

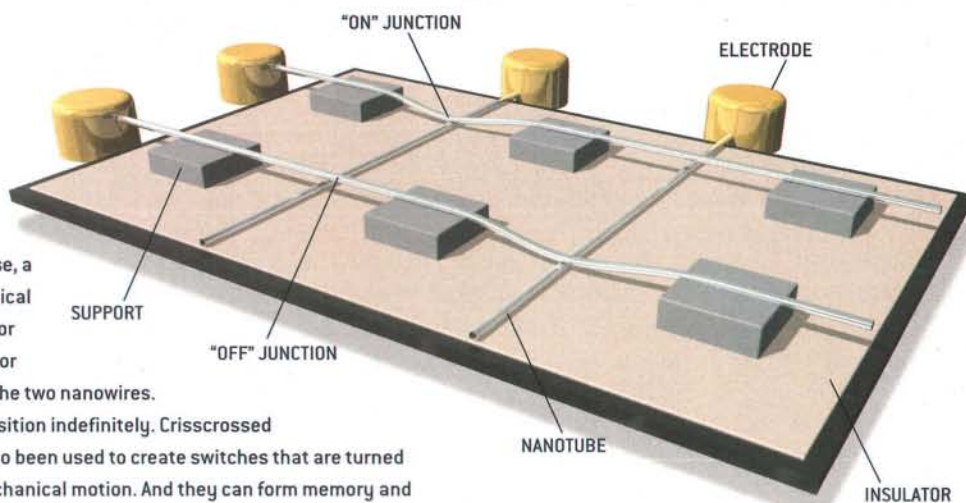
NANOWIRE ARRAY

CRISSCROSSING NANOWIRES

neatly solves a major problem in molecular-scale electronics:

How do you connect wires to components such as transistors or diodes? The wires do double duty, serving both as wires and as components. Each junction is a component, in this case, a miniature relay—an electromechanical switch that is either on (*touching*) or off (*separated*). To flip a switch on or off, you apply a certain voltage to the two nanowires.

The switch will then stay in that position indefinitely. Crisscrossed semiconductor nanowires have also been used to create switches that are turned on and off electrically, without mechanical motion. And they can form memory and logic arrays—key steps toward the assembly of a nanocomputer.



sued by researchers at U.C.L.A. and Hewlett-Packard, and it resembles the magnetic-core memory that was common in computers of the 1950s and 1960s.

Law of Large Numbers

TO OVERCOME the unreliability of individual nanodevices, we may rely on sheer numbers—the gizmos are so cheap that plenty of spares are always available. Researchers who work on defect tolerance have shown that computing is possible even if many of the components fail, although identifying and mapping the defects can be slow and time-consuming. Ultimately we hope to partition the enormous arrays into subarrays whose reliability can be easily monitored. The optimum size of these subarrays will depend on the defect levels typically present in molecular and nanoscale devices.

Another significant hurdle faced by nanoelectronics is “bootstrapping.” How do engineers get the circuit to do what they want it to? In microelectronics, circuit designers work like architects: they prepare a blueprint of a circuit, and a fabrication plant builds it. In nanoelectronics, designers will have to work like computer programmers. A fabrication plant will create a raw nanocircuit—billions on billions of devices and wires whose functioning is rather limited. From the outside, it will look like a lump of material with a handful of wires sticking out. Using those few wires, engineers will somehow have to configure those billions of devices. Challenges such as this are what keeps me tremendously excited about the field as a whole.

Even before we solve these problems, nanodevices may have useful applications. For example, semiconducting carbon nanotubes have been used by Hongjie Dai’s group at Stanford University to detect gas molecules, and Yi Cui in my group has used semiconductor nanowires as ultrasensitive detectors for a wide

range of biological compounds. In our work at Harvard, we have converted nanowire field-effect transistors into sensors by modifying their surfaces with molecular receptors. This technology has the potential of detecting single molecules using only a voltmeter from a hardware store. The small size and sensitivity of nanowires also make possible the assembly of extremely powerful sensors that could, for instance, sequence the entire human genome on a single chip and serve in minimally invasive medical devices. In the nearer term, we could see hybrids of micro and nano: silicon with a nano core—perhaps a high-density computer memory that retains its contents forever.

Although substantial work remains before nanoelectronics makes its way into computers, this goal now seems less hazy than it was even a year ago. As we gain confidence, we will learn not just to shrink digital microelectronics but also to go where no digital circuit has gone before. Nanoscale devices that exhibit quantum phenomena, for example, could be exploited in quantum encryption and quantum computing. The richness of the nanoworld will change the macroworld.

MORE TO EXPLORE

The author’s Web site: cmiiris.harvard.edu

The Avouris group: www.research.ibm.com/nanoscience

The Dai group: www-chem.stanford.edu/group/dai

The DeHon group: www.cs.caltech.edu/~andre

The Dekker group: www.mb.tn.tudelft.nl/user/dekker

The McEuen group:

www.lassp.cornell.edu/lassp_data/mceuen/homepage/welcome.html

The Penn State team: stm1.chem.psu.edu

The Rice/Yale team: www.jmtour.com and www.eng.yale.edu/reedlab

The Smith group: www.chem.wisc.edu/~smith

The U.C.L.A./Hewlett-Packard team: www.chem.ucla.edu/~schung/hgrp and www.hpl.hp.com/research/qsr/staff/kuekes.html

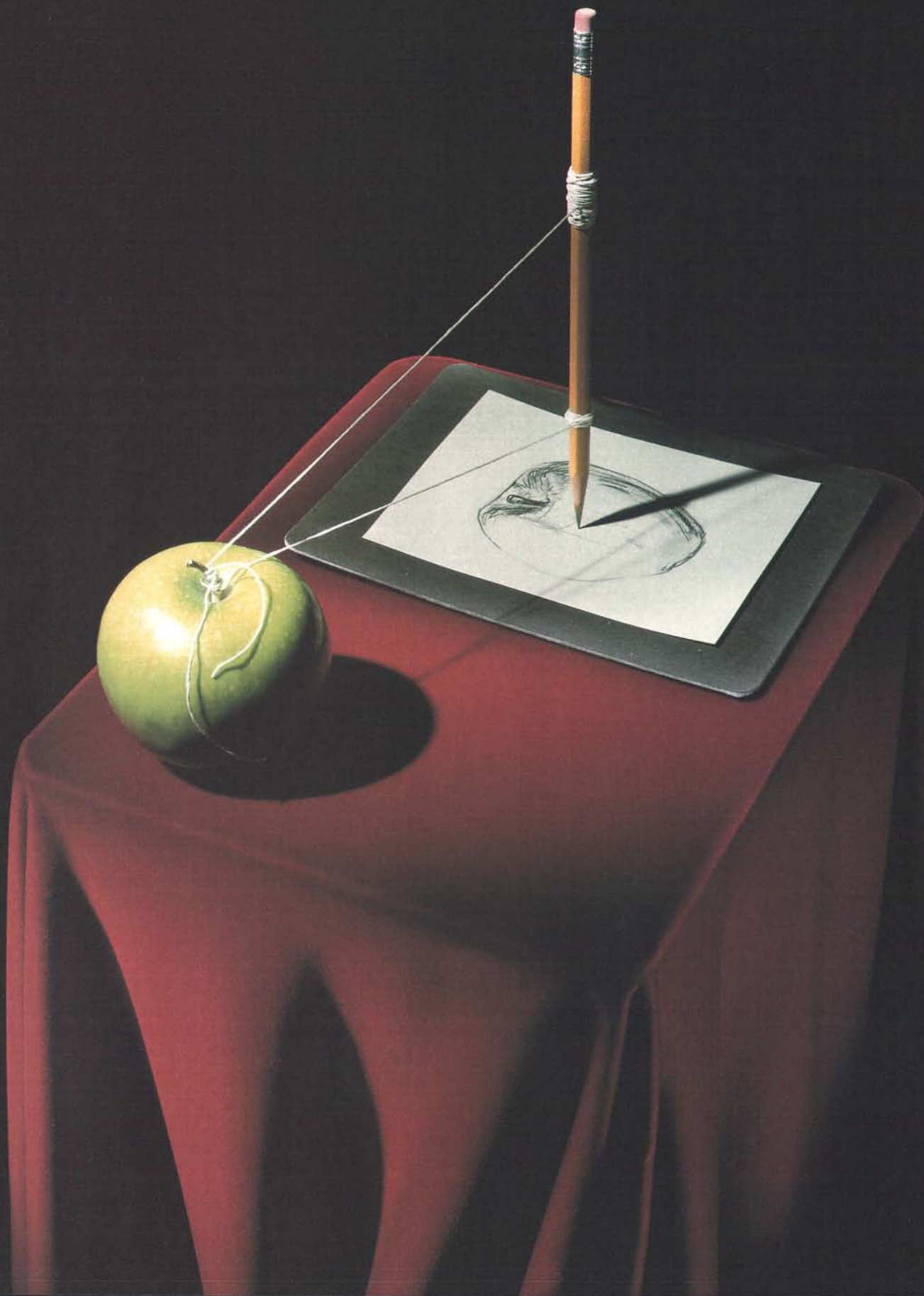
SCIENTIFIC AMERICAN.com

Everything in science & technology...
Plug yourself in.

- ▶ Today's News
- ▶ Ask the Experts
- ▶ Daily Trivia
- ▶ Quick Poll
- ▶ Scientific American Digital
- ▶ Careers
- ▶ Subscriptions
- ▶ Bookstore
- ▶ Shop

and more...

▶ ▶ ▶ WWW.ScientificAmerican.com



Go Forth and Replicate

Birds do it, bees do it,
but could machines do it?
New computer simulations
suggest that the answer is yes

Apples beget apples, but can machines beget machines? Today it takes an elaborate manufacturing apparatus to build even a simple machine. Could we endow an artificial device with the ability to multiply on its own? Self-replication has long been considered one of the fundamental properties separating the living from the nonliving. Historically our limited understanding of how biological reproduction works has given it an aura of mystery and made it seem unlikely that it would ever be done by a man-made object. It is reported that when René Descartes averred to Queen Christina of Sweden that animals were just another form of mechanical automata, Her Majesty pointed to a clock and said, "See to it that it produces offspring."

The problem of machine self-replication moved from philosophy into the realm of science and engineering in the late 1940s with the work of eminent mathematician and physicist John von Neumann. Some researchers have actually constructed physical replicators. Forty years ago, for example, geneticist Lionel Penrose and his son, Roger (the famous physicist), built small assemblies of plywood that exhibited a simple form of self-replication [see "Self-Reproducing Machines," by Lionel

Penrose; *SCIENTIFIC AMERICAN*, June 1959]. But self-replication has proved to be so difficult that most researchers study it with the conceptual tool that von Neumann developed: two-dimensional cellular automata.

Implemented on a computer, cellular automata can simulate a huge variety of self-replicators in what amount to austere universes with different laws of physics from our own. Such models free researchers from having to worry about logistical issues such as energy and physical construction so that they can focus on the fundamental questions of information flow. How is a living being able to replicate unaided, whereas mechanical objects must be constructed by humans? How does replication at the level of an organism emerge from the numerous interactions in tissues, cells and molecules? How did Darwinian evolution give rise to self-replicating organisms?

The emerging answers have inspired the development of self-repairing silicon chips [see box on page 46] and autocatalyzing molecules [see "Synthetic Self-Replicating Molecules," by Julius Rebek, Jr.; *SCIENTIFIC AMERICAN*, July 1994]. And this may be just the beginning. Researchers in the field of nanotechnology have long proposed that self-replication will be crucial to manu-

By Moshe Sipper and James A. Reggia

Photoillustrations by David Emmite

facturing molecular-scale machines, and proponents of space exploration see a macroscopic version of the process as a way to colonize planets using in situ materials. Recent advances have given credence to these futuristic-sounding ideas. As with other scientific disciplines, including genetics, nuclear energy and chemistry, those of us who study self-replication face the twofold challenge of creating replicating machines and avoiding dystopian pre-

scription could be used in two distinct ways: first, as the instructions whose interpretation leads to the construction of an identical copy of the device; next, as data to be copied, uninterpreted, and attached to the newly created child so that it too possesses the ability to self-replicate. With this two-step process, the self-description need not contain a description of itself. In the architectural analogy, the blueprint would include a plan for building a pho-

the cellular-automata world. All decisions and actions take place locally; cells do not know directly what is happening outside their immediate neighborhood.

The apparent simplicity of cellular automata is deceptive; it does not imply ease of design or poverty of behavior. The most famous automata, John Horton Conway's Game of Life, produces amazingly intricate patterns. Many questions about the dynamic behavior of cellular

Her Majesty pointed to a clock and said, "See to it that it produces offspring."

dictions of devices running amok. The knowledge we gain will help us separate good technologies from destructive ones.

Playing Life

SCIENCE-FICTION STORIES often depict cybernetic self-replication as a natural development of current technology, but they gloss over the profound problem it poses: how to avoid an infinite regress. A system might try to build a clone using a blueprint—that is, a self-description. Yet the self-description is part of the machine, is it not? If so, what describes the description? And what describes the description of the description? Self-replication in this case would be like asking an architect to make a perfect blueprint of his or her own studio. The blueprint would have to contain a miniature version of the blueprint, which would contain a miniature version of the blueprint and so on. Without this information, a construction crew would be unable to re-create the studio fully; there would be a blank space where the blueprint had been.

Von Neumann's great insight was an explanation of how to break out of the infinite regress. He realized that the self-de-

tocopy machine. Once the new studio and the photocopier were built, the construction crew would simply run off a copy of the blueprint and put it into the new studio.

Living cells use their self-description, which biologists call the genotype, in exactly these two ways: transcription (DNA is copied mostly uninterpreted to form mRNA) and translation (mRNA is interpreted to build proteins). Von Neumann made this transcription-translation distinction several years before molecular biologists did, and his work has been crucial in understanding self-replication in nature.

To prove these ideas, von Neumann and mathematician Stanislaw M. Ulam came up with the idea of cellular automata. A cellular-automata simulation involves a chessboardlike grid of squares, or cells, each of which is either empty or occupied by one of several possible components. At discrete intervals of time, each cell looks at itself and its neighbors and decides whether to metamorphose into a different component. In making this decision, the cell follows relatively simple rules, which are the same for all cells. These rules constitute the basic physics of

automata are formally unsolvable. To see how a pattern will unfold, you need to simulate it fully [see *Mathematical Games*, by Martin Gardner; *SCIENTIFIC AMERICAN*, October 1970 and February 1971; and "The Ultimate in Anty-Particles," by Ian Stewart, July 1994]. In its own way, a cellular-automata model can be just as complex as the real world.

Copy Machines

WITHIN CELLULAR AUTOMATA, self-replication occurs when a group of components—a "machine"—goes through a sequence of steps to construct a nearby duplicate of itself. Von Neumann's machine was based on a universal constructor, a machine that, given the appropriate instructions, could create any pattern. The constructor consisted of numerous types of components spread over tens of thousands of cells and required a book-length manuscript to be specified. It has still not been simulated in its entirety, let alone actually built, on account of its complexity. A constructor would be even more complicated in the Game of Life because the functions performed by single cells in von Neumann's model—such as transmission of signals and generation of new components—have to be performed by composite structures in Life.

Going to the other extreme, it is easy to find trivial examples of self-replication. For example, suppose a cellular automata has only one type of component, labeled +, and that each cell follows only a single rule: if exactly one of the four neighboring

cells contains a +, then the cell becomes a +; otherwise it becomes vacant. With this rule, a single + grows into four more +'s, each of which grows likewise, and so forth.

Such weedlike proliferation does not shed much light on the principles of replication, because there is no significant machine. Of course, that invites the question of how you would tell a "significant" machine from a trivially prolific automata. No one has yet devised a satisfactory answer. What is clear, however, is that the replicating structure must in some sense be complex. For example, it must consist of multiple, diverse components whose interactions collectively bring about replication—the proverbial "whole must be greater than the sum of the parts." The existence of multiple distinct components permits a self-description to be stored within the replicating structure.

In the years since von Neumann's seminal work, many researchers have probed the domain between the complex and the trivial, developing replicators that require fewer components, less space or simpler rules. A major step forward was taken in 1984 when Christopher G. Langton, then at the University of Michigan, observed that looplike storage devices—which had formed modules of earlier self-replicating machines—could be programmed to replicate on their own. These devices typically consist of two pieces: the loop itself, which is a string of components that circulate around a rectangle, and a construction arm, which protrudes from a corner of the rectangle into the surrounding space. The circulating components constitute a recipe for the loop—for example, "go three squares ahead, then turn left." When this recipe reaches the construction arm, the automata rules make a copy of it. One copy continues around the loop; the other goes down the arm, where it is interpreted as instructions.

By giving up the requirement of universal construction, which was central to von Neumann's approach, Langton showed that a replicator could be constructed from just seven unique components occupying only 86 cells. Even smaller and simpler self-replicating loops have been devised by one of us (Reggia) and our colleagues [see box on next page]. Be-

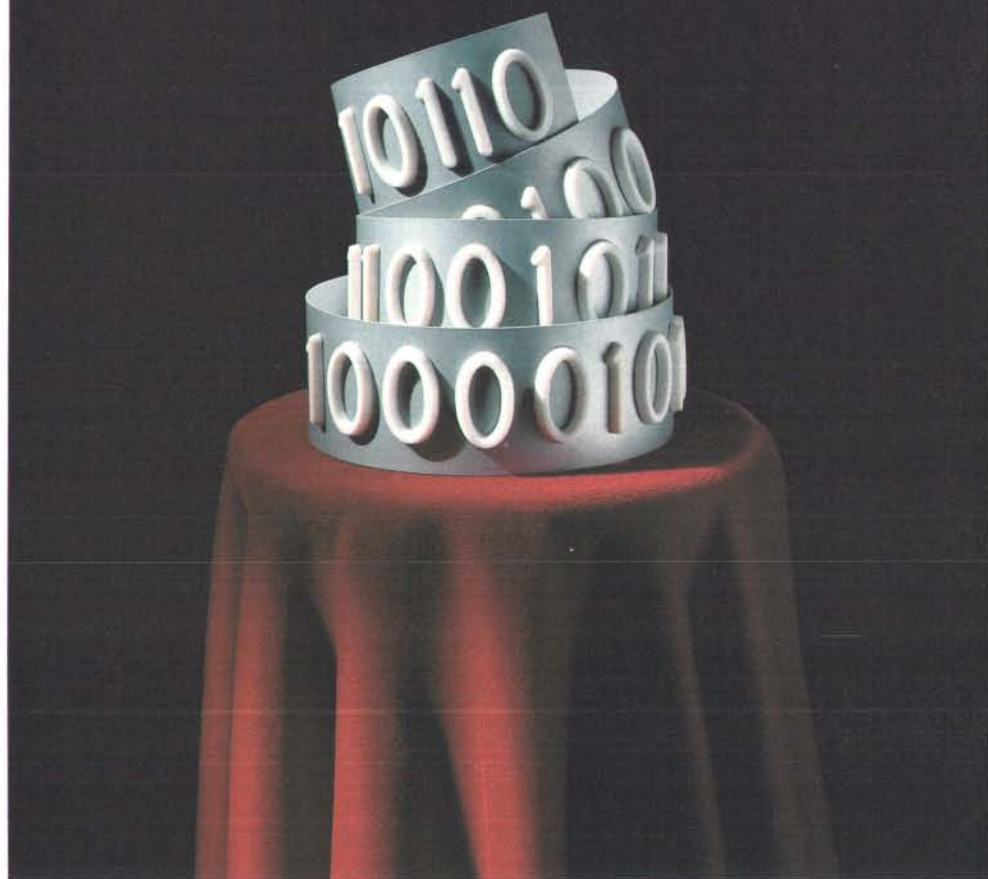
cause they have multiple interacting components and include a self-description, they are not trivial. Intriguingly, asymmetry plays an unexpected role: the rules governing replication are often simpler when the components are not rotationally symmetric than when they are.

Emergent Replication

ALL THESE SELF-REPLICATING structures have been designed through ingenuity and much trial and error. This process is arduous and often frustrating; a small change to one of the rules results in an entirely different global behavior, most likely the disintegration of the structure in question. But recent work has gone beyond the direct-design approach. Instead of tailoring the rules to suit a par-

ticular type of structure, researchers have experimented with various sets of rules, filled the cellular-automata grid with a "primordial soup" of randomly selected components and checked whether self-replicators emerged spontaneously.

In 1997 Hui-Hsien Chou, now at Iowa State University, and Reggia noticed that as long as the initial density of the free-floating components was above a certain threshold, small self-replicating loops reliably appeared. Loops that collided underwent annihilation, so there was an ongoing process of death as well as birth. Over time, loops proliferated, grew in size and evolved through mutations triggered by debris from past collisions. Although the automata rules were deterministic, these mutations were effectively random,



because the system was complex and the components started in random locations.

Such loops are intended as abstract machines and not as simulacra of anything biological, but it is interesting to compare them with biomolecular structures. A loop loosely resembles circular DNA in bacteria, and the construction arm acts as the enzyme that catalyzes DNA replication. More important, replicating loops illustrate how complex global behaviors can arise from simple local in-

teractions. For example, components move around a loop even though the rules say nothing about movement; what is actually happening is that individual cells are coming alive, dying or metamorphosing in such a way that a pattern is eliminated from one position and reconstructed elsewhere—a process that we perceive as motion. In short, cellular automata act locally but appear to think globally. Much the same is true of molecular biology.

In a recent computational experiment,

Jason Lohn, now at the NASA Ames Research Center, and Reggia experimented not with different structures but with different sets of rules. Starting with an arbitrary block of four components, they found they could determine a set of rules that made the block self-replicate. They discovered these rules via a genetic algorithm, an automated process that simulates Darwinian evolution.

The most challenging aspect of this work was the definition of the so-called

BUILD YOUR OWN REPLICATOR

SIMULATING A SMALL self-replicating loop using an ordinary chess set is a good way to get an intuitive sense of how these systems work. This particular cellular-automata model has four different types of components: pawns, knights, bishops and rooks. The machine initially comprises four pawns, a knight and a bishop. It has two parts: the loop itself, which consists of a two-by-two square, and a construction arm, which sticks out to the right.

The knight and bishop represent the self-description: the knight, whose orientation is significant, determines which direction to grow, while the bishop tags along and determines how long the side of the loop should be. The pawns are fillers that define the rest of the shape of the loop, and the rook is a transient signal to guide the growth of a new construction arm.

As time progresses, the knight and bishop circulate counterclockwise around the loop. Whenever they encounter the arm, one copy goes out the arm while the original continues around the loop.

HOW TO PLAY: You will need two chessboards: one to represent the current configuration, the other to show the next configuration. For each round, look at each square of the current configuration, consult the rules and place the appropriate piece in the corresponding square on the other board. Each piece metamorphoses depending on its identity and that of the four squares immediately to the left, to the right, above and below. When you have reviewed each square and set up the next configuration, the round is over. Clear the first board and repeat. Because the rules are complicated, it takes a bit of patience at first. You can also view the simulation at lslwww.epfl.ch/chess

The direction in which a knight faces is significant. In the drawings here, we use standard chess conventions to indicate the orientation of the knight: the horse's muzzle points forward. If no rule explicitly applies, the contents of the square stay the same. Squares on the edge should be treated as if they have adjacent empty squares off the board. —M.S. and J.A.R.

STAGES OF REPLICATION



INITIALLY, the self-description, or "genome"—a knight followed by a bishop—is poised at the start of the construction arm.



1 The knight and bishop move counterclockwise around the loop. A clone of the knight heads out the arm.



2 The original knight-bishop pair continues to circulate. The bishop is cloned and follows the new knight out the arm.



3 The knight triggers the formation of two corners of the child loop. The bishop tags along, completing the gene transfer.



4 The knight forges the remaining corner of the child loop. The loops are connected by the construction arm and a knight-errant.

fitness function—the criteria by which sets of rules were judged, thus separating good solutions from bad ones and driving the evolutionary process toward rule sets that facilitated replication. You cannot simply assign high fitness to those sets of rules that cause a structure to replicate, because none of the initial rule sets is likely to allow for replication. The solution was to devise a fitness function composed of a weighted sum of three measures: a growth measure (the extent to which

each component type generates an increasing supply of that component), a relative position measure (the extent to which neighboring components stay together) and a replicant measure (a function of the number of actual replicators present). With the right fitness function, evolution can turn rule sets that are sterile into ones that are fecund; the process usually takes 150 or so generations.

Self-replicating structures discovered in this fashion work in a fundamentally

different way than self-replicating loops do. For example, they move and deposit copies along the way—unlike replicating loops, which are essentially static. And although these newly discovered replicators consist of multiple, locally interacting components, they do not have an identifiable self-description—there is no obvious genome. The ability to replicate without a self-description may be relevant to questions about how the earliest biological

Continued on page 49

KNIGHT



IF THERE is a bishop just behind or to the left of the knight, replace the knight with another bishop.



OTHERWISE, if at least one of the neighboring squares is occupied, remove the knight and leave the square empty.

PAWN

IF THERE is a neighboring knight, replace the pawn with a knight with a certain orientation, as follows:



IF A NEIGHBORING knight is facing away from the pawn, the new knight faces the opposite way.



OTHERWISE, if there is exactly one neighboring pawn, the new knight faces that pawn.



OTHERWISE the new knight faces in the same direction as the neighboring knight.

BISHOP OR ROOK



REPLACE IT with a pawn.



EMPTY SQUARE



IF THERE are two neighboring knights and either faces the empty square, fill the square with a rook.



IF THERE is only one neighboring knight and it faces the square, fill the square with a knight rotated 90 degrees counterclockwise.



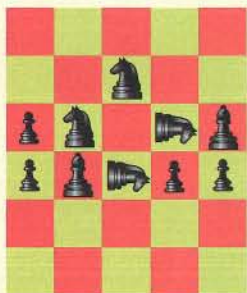
IF THERE is a neighboring knight and its left side faces the square, and the other neighbors are empty, fill the square with a pawn.



IF THERE is a neighboring rook, and the other neighbors are empty, fill the square with a pawn.



IF THERE are three neighboring pawns, fill the square with a knight facing the fourth, empty neighbor.



5 The knight-errant moves up to endow the parent with a new arm. A similar process, one step delayed, begins for the child loop.



6 The knight-errant, together with the original knight-bishop pair, conjures up a rook. Meanwhile the old arm is erased.



7 The rook kills the knight and generates the new, upward arm. Another rook prepares to do the same for the child.



8 At last the two loops are separate and whole. The self-descriptions continue to circulate, but otherwise all is calm.



9 The parent prepares to give birth again. In the following step, the child too will begin to replicate.

ROBOT, HEAL THYSELF

Computers that fix themselves are the first application of artificial self-replication

LAUSANNE, SWITZERLAND—Not many researchers encourage the wanton destruction of equipment in their labs. Daniel Mange, however, likes it when visitors walk up to one of his inventions and press the button marked KILL. The lights on the panel go out; a small box full of circuitry is toast. Early in May his team unveiled its latest contraption at a science festival here—a wall-size digital clock whose components you can zap at will—and told the public: Give it your best shot. See if you can crash the system.

The goal of Mange and his team is to instill electronic circuits with the ability to take a lickin' and keep on tickin'—just like living things. Flesh-and-blood creatures might not be so good at calculating π to the millionth digit, but they can get through the day without someone pressing Ctrl-Alt-Del. Combining the precision of digital hardware with the resilience of biological wetware is a leading challenge for modern electronics.

Electronics engineers have been working on fault-tolerant circuits ever since there were electronics engineers [see "Redundancy in Computers," by William H. Pierce; SCIENTIFIC AMERICAN, February 1964]. Computer modems would still be dribbling data at 1200 baud if it weren't for error detection and correction. In many applications, simple quality-control checks, such as extra data bits, suffice. More complex systems provide entire backup computers. The space shuttle, for example, has five processors. Four of them perform the same calculations; the fifth checks whether they agree and pulls the plug on any dissenter.

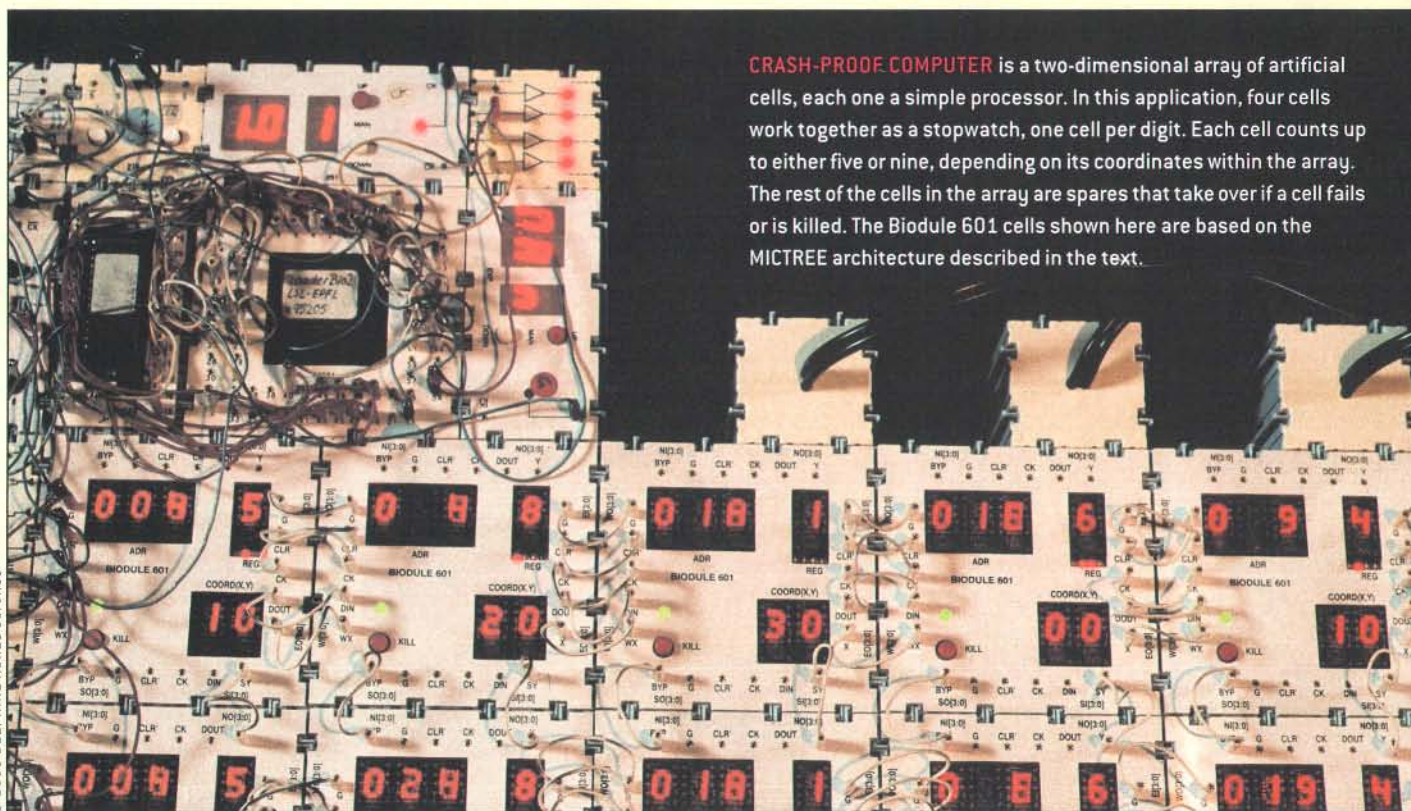
The problem with these systems, though, is that they rely on centralized control. What if that control unit goes bad?

Nature has solved that problem through radical decentralization. Cells in the body are all basically identical; each takes on a specialized task, performs it autonomously and, in the event of infection or failure, commits hara-kiri so that its tasks can be taken up by new cells. These are the attributes that Mange, a professor at the Swiss Federal Institute of Technology here, and others have sought since 1993 to emulate in circuitry, as part of the "Embryonics" (embryonic electronics) project.

One of their earlier inventions, the MICTREE (microinstruction tree) artificial cell, consisted of a simple processor and four bits of data storage. The cell is contained in a plastic box roughly the size of a pack of Post-its. Electrical contacts run along the sides so that cells can be snapped together like Legos. As in cellular automata, the models used to study the theory of self-replication, the MICTREE cells are connected only to their immediate neighbors. The communication burden on each cell is thus independent of the total number of cells. The system, in other words, is easily scalable—unlike many parallel-computing architectures.

Cells follow the instructions in their "genome," a program written in a subset of the Pascal computer language. Like their biological antecedents, the cells all contain the exact same genome and execute part of it based on their position within the array, which each cell calculates relative to its neighbors. Waste-

CRASH-PROOF COMPUTER is a two-dimensional array of artificial cells, each one a simple processor. In this application, four cells work together as a stopwatch, one cell per digit. Each cell counts up to either five or nine, depending on its coordinates within the array. The rest of the cells in the array are spares that take over if a cell fails or is killed. The Biodule 601 cells shown here are based on the MICTREE architecture described in the text.



ful though it may seem, this redundancy allows the array to withstand the loss of any cell. Whenever someone presses the KILL button on a cell, that cell shuts down, and its left and right neighbors become directly connected. The right neighbor recalculates its position and starts executing the deceased's program. Its tasks, in turn, are taken up by the next cell to the right, and so on, until a cell designated as a spare is pressed into service.

Writing programs for any parallel processor is tricky, but the MICTREE array requires an especially unconventional approach. Instead of giving explicit instructions, the programmer must devise simple rules out of which the desired function will emerge. Being Swiss, Mange demonstrates by building a superreliable stopwatch. Displaying minutes and seconds requires four cells in a row, one for each digit. The genome allows for two cell types: a counter from zero to nine and a counter from zero to five. An oscillator feeds one pulse per second into the rightmost cell. After 10 pulses, this cell cycles back to zero and sends a pulse to the cell on its left, and so on down the line. The watch takes up part of an array of 12 cells; when you kill one, the clock transplants itself one cell over and carries on. Obviously, though, there is a limit to its resilience: the whole thing will fail after, at most, eight kills.

The prototype MICTREE cells are hardwired, so their processing power cannot be tailored to a specific application. In a finished product, cells would instead be implemented on a field-programmable gate array, a grid of electronic components that can be reconfigured on the fly [see "Configurable Computing," by John Villasenor and William H. Mangione-Smith; *SCIENTIFIC AMERICAN*, June 1997]. Mange's team is now custom-designing a gate array,

known as MUXTREE (multiplexer tree), that is optimized for artificial cells. In the biological metaphor, the components of this array are the "molecules" that constitute a cell. Each consists of a logic gate, a data bit and a string of configuration bits that determines the function of this gate.

Building a cell out of such molecules offers not only flexibility but also extra endurance. Each molecule contains two copies of the gate and three of the storage bit. If the two gates ever give different results, the molecule kills itself for the greater good of the cell. As a last gasp, the molecule sends its data bit (preserved by the triplicate storage) and configuration to its right neighbor, which does the same, and the process continues until the rightmost molecule transfers its data to a spare. This second level of fault tolerance prevents a single error from wiping out an entire cell.

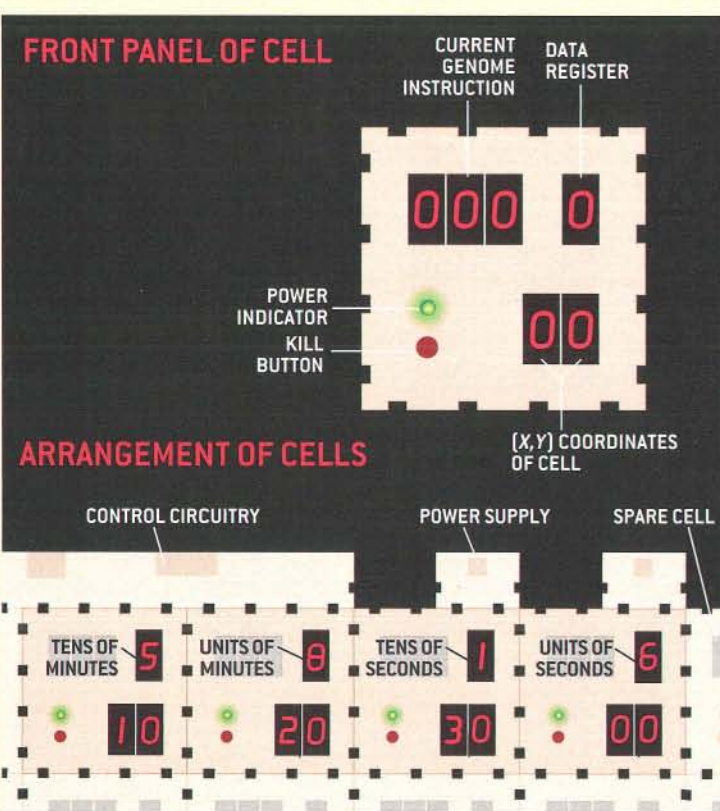
A total of 2,000 molecules, divided into four 20-by-25 cells, make up the BioWall—the giant digital clock that Mange's team has just put on display. Each molecule is enclosed in a small box and includes a KILL button and an LED display. Some molecules are configured to perform computations; others serve as pixels in the clock display. Making liberal use of the KILL buttons, I did my utmost to crash the system, something I'm usually quite good at. But the plucky clock just wouldn't submit. The clock display did start to look funny—numerals bent over as their pixels shifted to the right—but at least it was still legible, unlike most faulty electronic signs.

That said, the system did suffer from display glitches, which Mange attributed mainly to timing problems. Although the processing power is decentralized, the cells still rely on a central oscillator to coordinate their communications; sometimes they fall out of sync. Another Embryonics team, led by Andy Tyrrell of the University of York in England, has been studying making the cells asynchronous, like their biological counterparts. Cells would generate handshaking signals to orchestrate data transfers. The present system is also unable to catch certain types of error, including damaged configuration strings. Tyrrell's team has proposed adding watchdog molecules—an immune system—that would monitor the configurations (and one another) for defects.

Although these systems demand an awful lot of overhead, so do other fault-tolerance technologies. "While Embryonics appears to be heavy on redundancy, it actually is not that bad when compared to other systems," Tyrrell argues. Moreover, MUXTREE should be easier to scale down to the nano level; the "molecules" are simple enough to really be molecules. Says Mange, "We are preparing for the situation where electronics will be at the same scale as biology."

On a philosophical level, Embryonics comes very close to the dream of building a self-replicating machine. It may not be quite as dramatic as a robot that can go down to Radio Shack, pull parts off the racks, and take them home to resolder a connection or build a loving mate. But the effect is much the same. Letting machines determine their own destiny—whether reconfiguring themselves on a silicon chip or reprogramming themselves using a neural network or genetic algorithm—sounds scary, but perhaps we should be gratified that machines are becoming more like us: imperfect, fallible but stubbornly resourceful.

—George Musser, *imperfect but resourceful* staff editor and writer





Continued from page 45

replicators originated. In a sense, researchers are seeing a continuum between nonliving and living structures.

Many researchers have tried other computational models besides the traditional cellular automata. In asynchronous cellular automata, cells are not updated in concert; in nonuniform cellular automata, the rules can vary from cell to cell. Another approach altogether is Core War [see Computer Recreations, by A. K. Dewdney; SCIENTIFIC AMERICAN, May 1984] and its successors, such as ecologist Thomas S. Ray's Tierra system. In these

nents, one for the program and the other for data. The loops can execute an arbitrary program in addition to self-replicating. In a sense, they are as complex as the computer that simulates them. Their main limitation is that the program is copied unchanged from parent to child, so that all loops carry out the same set of instructions.

In 1998 Chou and Reggia swept away this limitation. They showed how self-replicating loops carrying distinct information, rather than a cloned program, can be used to solve a problem known as satisfiability. The loops can be used to determine whether the variables in a logical ex-

pressing a parallel computer from either transistors or chemicals [see "Computing with DNA," by Leonard M. Adleman; SCIENTIFIC AMERICAN, August 1998].

In 1980 a NASA team led by Robert Freitas, Jr., proposed planting a factory on the moon that would replicate itself, using local lunar materials, to populate a large area exponentially. Indeed, a similar probe could colonize the entire galaxy, as physicist Frank J. Tipler of Tulane University has argued. In the nearer term, computer scientists and engineers have experimented with the automated design of robots [see "Dawn of a New Species?" by George

In a sense, researchers are seeing a continuum between nonliving and living structures.

simulations the "organisms" are computer programs that vie for processor time and memory. Ray has observed the emergence of "parasites" that co-opt the self-replication code of other organisms.

Getting Real

SO WHAT GOOD are these machines? Von Neumann's universal constructor can compute in addition to replicating, but it is an impractical beast. A major advance has been the development of simple yet useful replicators. In 1995 Gianluca Tempesti of the Swiss Federal Institute of Technology in Lausanne simplified the loop self-description so it could be interlaced with a small program—in this case, one that would spell the acronym of his lab, "LSL." His insight was to create automata rules that allow loops to replicate in two stages. First the loop, like Langton's loop, makes a copy of itself. Once finished, the daughter loop sends a signal back to its parent, at which point the parent sends the instructions for writing out the letters.

Drawing letters was just a demonstration. The following year Jean-Yves Perrier, Jacques Zahnd and one of us (Sipper) designed a self-replicating loop with universal computational capabilities—that is, with the computational power of a universal Turing machine, a highly simplified but fully capable computer. This loop has two "tapes," or long strings of compo-

pression can be assigned values such that the entire expression evaluates to "true." This problem is NP-complete—in other words, it belongs to the family of nasty puzzles, including the famous traveling-salesman problem, for which there is no known efficient solution. In Chou and Reggia's cellular-automata universe, each replicator received a different partial solution. During replication, the solutions mutated, and replicators with promising solutions were allowed to proliferate while those with failed solutions died out.

Although various teams have created cellular automata in electronic hardware, such systems are probably too wasteful for practical applications; automata were never really intended to be implemented directly. Their purpose is to illuminate the underlying principles of replication and, by doing so, inspire more concrete efforts. The loops provide a new paradigm for de-

Musser; SCIENTIFIC AMERICAN, November 2000]. Although these systems are not truly self-replicating—the offspring are much simpler than the parent—they are a first step toward fulfilling the queen of Sweden's request.

Should physical self-replicating machines become practical, they and related technologies will raise difficult issues, including the *Terminator* film scenario in which artificial creatures outcompete natural ones. We prefer the more optimistic, and more probable, scenario that replicators will be harnessed to the benefit of humanity [see "Will Robots Inherit the Earth?" by Marvin Minsky; SCIENTIFIC AMERICAN, October 1994]. The key will be taking the advice of 14th-century English philosopher William of Ockham: *entia non sunt multiplicanda praeter necessitatem*—entities are not to be multiplied beyond necessity. ■

MORE TO EXPLORE

Simple Systems That Exhibit Self-Directed Replication. J. Reggia, S. Armentrout, H. Chou and Y. Peng in *Science*, Vol. 259, No. 5099, pages 1282–1287; February 26, 1993.

Emergence of Self-Replicating Structures in a Cellular Automata Space. H. Chou and J. Reggia in *Physica D*, Vol. 110, Nos. 3–4, pages 252–272; December 15, 1997.

Special Issue: Von Neumann's Legacy: On Self-Replication. Edited by M. Sipper, G. Tempesti, D. Mange and E. Sanchez in *Artificial Life*, Vol. 4, No. 3; Summer 1998.

Towards Robust Integrated Circuits: The Embryonics Approach. D. Mange, M. Sipper, A. Stauffer and G. Tempesti in *Proceedings of the IEEE*, Vol. 88, No. 4, pages 516–541; April 2000.

Moshe Sipper's Web page on artificial self-replication is at islwww.epfl.ch/~moshes/selfrep/

Animations of self-replicating loops can be found at necsi.org/postdocs/sayama/sdsr/java/

For John von Neumann's universal constructor, see alife.santafe.edu/alife/topics/jvn/jvn.html



The Do-It-Yourself Supercom

By William W. Hargrove,
Forrest M. Hoffman and
Thomas Sterling

Photographs by Kay Chernush

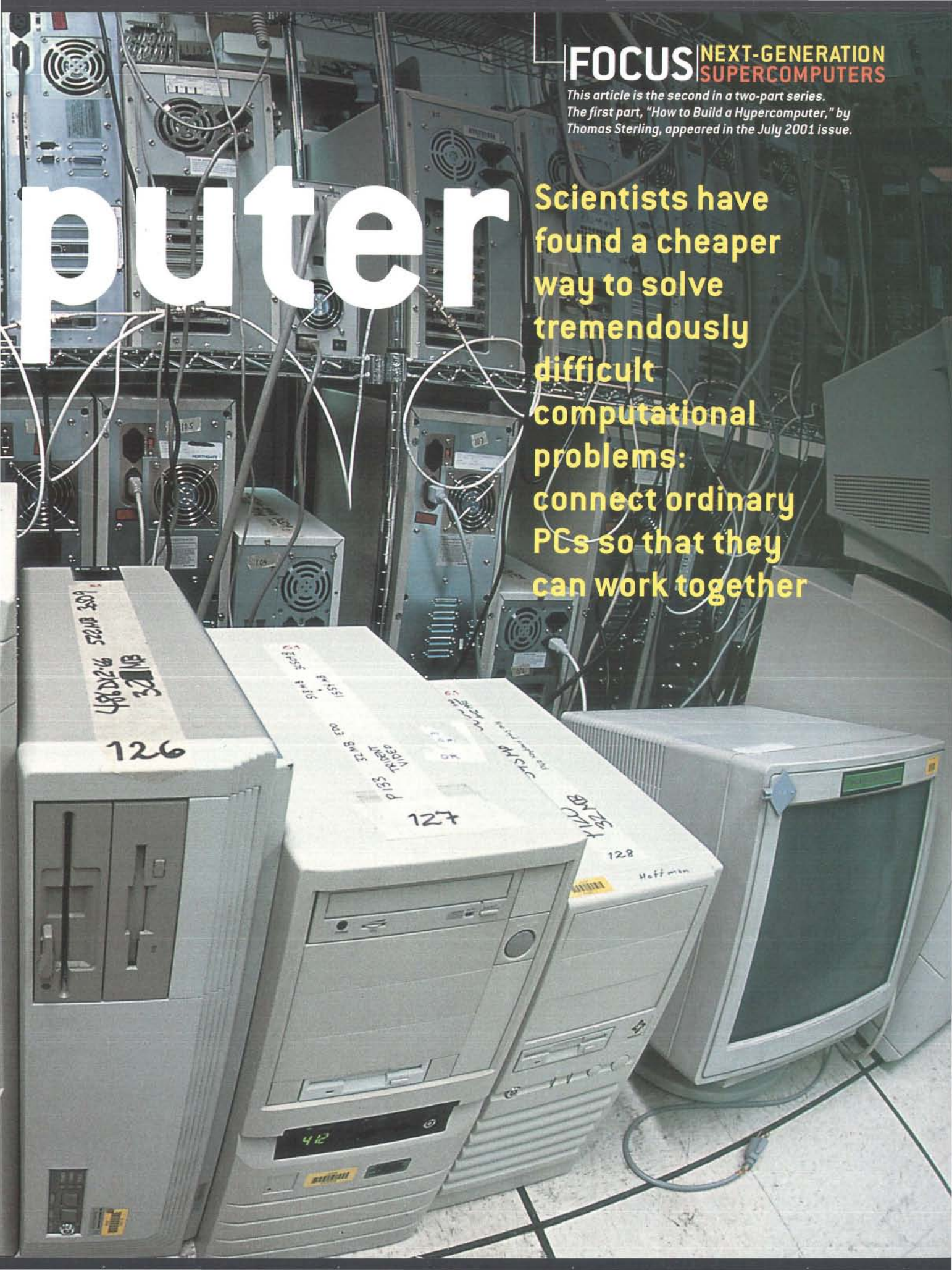
CLUSTER OF PCs at the
Oak Ridge National
Laboratory in Tennessee
has been dubbed the
Stone SouperComputer.

FOCUS | NEXT-GENERATION SUPERCOMPUTERS

This article is the second in a two-part series.
The first part, "How to Build a Hypercomputer," by
Thomas Sterling, appeared in the July 2001 issue.

puter

Scientists have
found a cheaper
way to solve
tremendously
difficult
computational
problems:
connect ordinary
PCs so that they
can work together



IN THE WELL-KNOWN STONE SOUP FABLE, a wandering soldier stops at a poor village and says he will make soup by boiling a cauldron of water containing only a shiny stone. The townspeople are skeptical at first but soon bring small offerings: a head of cabbage, a bunch of carrots, a bit of beef. In the end, the cauldron is filled with enough hearty soup to feed everyone. The moral: cooperation can produce significant achievements, even from meager, seemingly insignificant contributions.

Researchers are now using a similar cooperative strategy to build supercomputers, the powerful machines that can perform billions of calculations in a second. Most conventional supercomputers employ parallel processing: they contain arrays of ultrafast microprocessors that work in tandem to solve complex problems such as forecasting the weather or simulating a nuclear explosion. Made by IBM, Cray and other computer vendors, the machines typically cost tens of millions of dollars—far too much for a research team with a modest budget. So over the past few years, scientists at national laboratories and universities have learned how to construct their own supercomputers by linking inexpensive PCs and writing software that allows these ordinary computers to tackle extraordinary problems.

In 1996 two of us (Hargrove and Hoffman) encountered such a problem in our work at Oak Ridge National Laboratory (ORNL) in Tennessee. We were trying to draw a national map of ecoregions, which are defined by environmental conditions: all areas with the same climate, landforms and soil characteristics fall into the same ecoregion. To create a high-resolution map of the continental U.S., we divided the country into 7.8 million square cells, each with an area of one square kilometer. For each cell we had to consider as many as 25 variables, ranging from average monthly precipitation to the nitrogen content of the soil. A single PC or workstation could not accomplish the task. We needed a parallel-processing supercomputer—and one that we could afford!

Our solution was to construct a com-

puting cluster using obsolete PCs that ORNL would have otherwise discarded. Dubbed the Stone SouperComputer because it was built essentially at no cost, our cluster of PCs was powerful enough to produce ecoregion maps of unprecedented detail. Other research groups have devised even more capable clusters that rival the performance of the world's best supercomputers at a mere fraction of their cost. This advantageous price-to-performance ratio has already attracted the attention of some corporations, which plan to use the clusters for such complex tasks as deciphering the human genome. In fact, the cluster concept

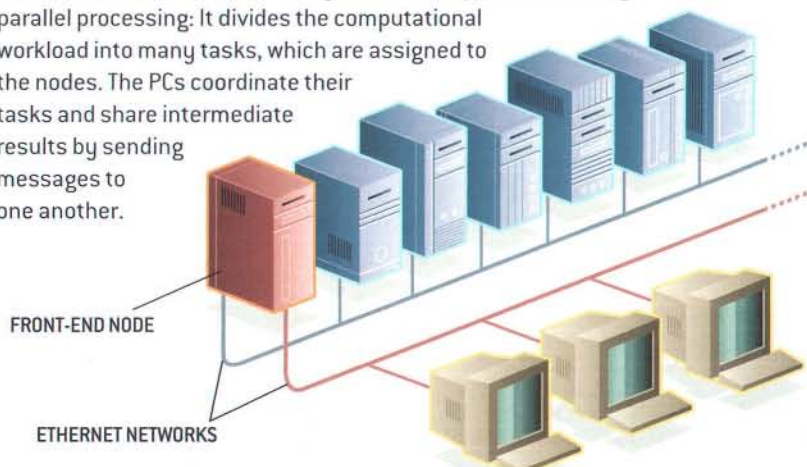
promises to revolutionize the computing field by offering tremendous processing power to any research group, school or business that wants it.

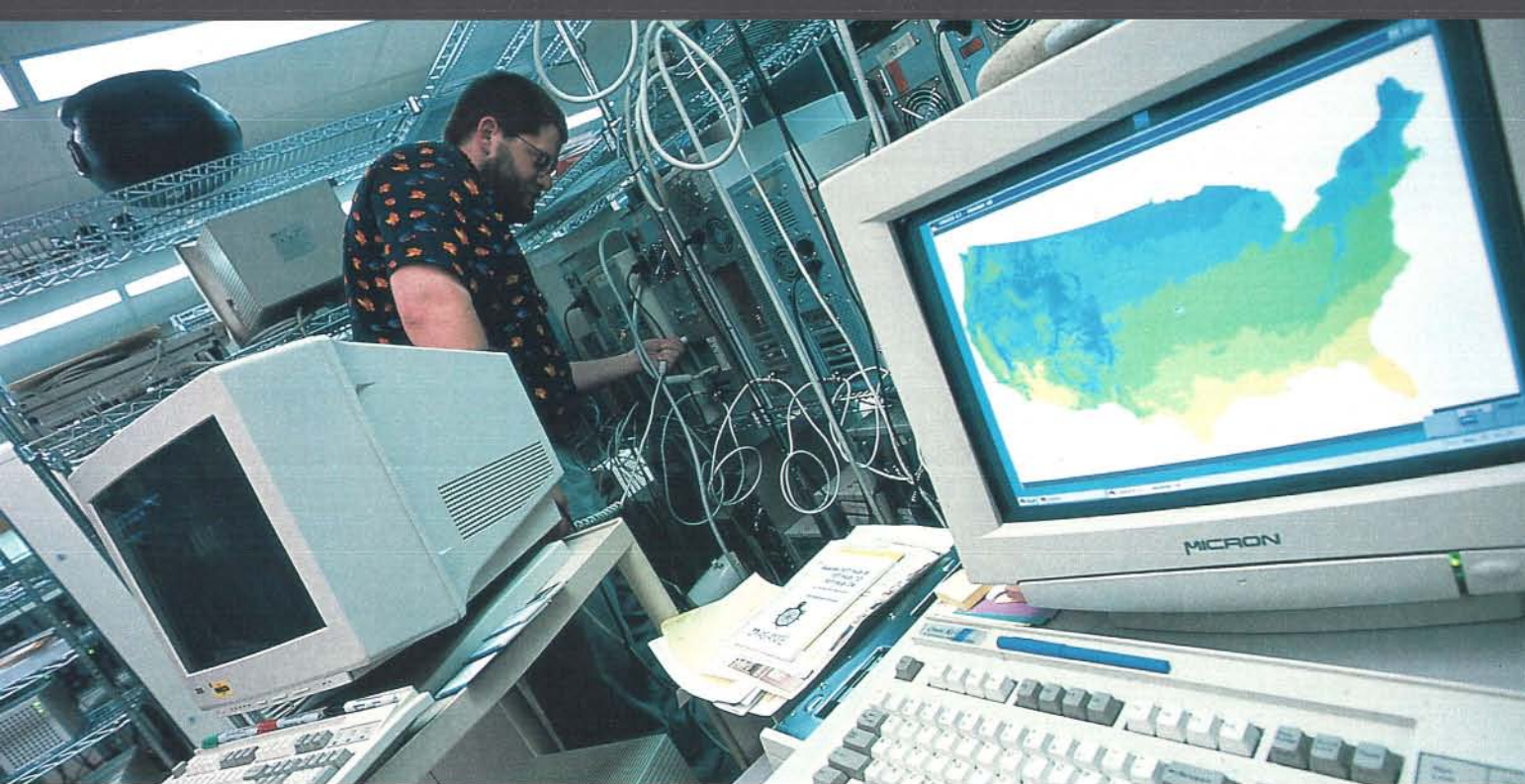
Beowulf and Grendel

THE NOTION OF LINKING computers together is not new. In the 1950s and 1960s the U.S. Air Force established a network of vacuum-tube computers called SAGE to guard against a Soviet nuclear attack. In the mid-1980s Digital Equipment Corporation coined the term "cluster" when it integrated its mid-range VAX minicomputers into larger systems. Networks of workstations—generally

A COMPUTING CLUSTER

The Stone SouperComputer at Oak Ridge National Laboratory consists of more than 130 PCs linked in a computing cluster. One of the machines serves as the front-end node for the cluster; it has two Ethernet cards, one for communicating with users and outside networks, and the other for talking with the rest of the nodes in the cluster. The system solves problems through parallel processing: It divides the computational workload into many tasks, which are assigned to the nodes. The PCs coordinate their tasks and share intermediate results by sending messages to one another.





less powerful than minicomputers but faster than PCs—soon became common at research institutions. By the early 1990s scientists began to consider building clusters of PCs, partly because their mass-produced microprocessors had become so inexpensive. What made the idea even more appealing was the falling cost of Ethernet, the dominant technology for connecting computers in local-area networks.

Advances in software also paved the way for PC clusters. In the 1980s Unix emerged as the dominant operating system for scientific and technical computing. Unfortunately, the operating systems for PCs lacked the power and flexibility of Unix. But in 1991 Finnish college student Linus Torvalds created Linux, a Unix-like operating system that ran on a PC. Torvalds made Linux available free of charge on the Internet, and soon hundreds of programmers began contributing improvements. Now wildly popular as an operating system for stand-alone computers, Linux is also ideal for clustered PCs.

The first PC cluster was born in 1994 at the NASA Goddard Space Flight Center. NASA had been searching for a cheaper way to solve the knotty computational problems typically encountered in earth and space science. The space agency needed a machine that could achieve one gigaflops—that is, perform a billion floating-

"CRASH CART" with a monitor and keyboard diagnoses problems with the Stone SouperComputer.

point operations per second. (A floating-point operation is equivalent to a simple calculation such as addition or multiplication.) At the time, however, commercial supercomputers with that level of performance cost about \$1 million, which was too expensive to be dedicated to a single group of researchers.

One of us (Sterling) decided to pursue the then radical concept of building a computing cluster from PCs. Sterling and his Goddard colleague Donald J. Becker connected 16 PCs, each containing an Intel 486 microprocessor, using Linux and a standard Ethernet network. For scientific applications, the PC cluster delivered sustained performance of 70 megaflops—that is, 70 million floating-point operations per second. Though modest by today's standards, this speed was not much lower than that of some smaller commercial supercomputers available at the time. And the cluster was built for only \$40,000, or about one tenth the price of a comparable commercial machine in 1994.

NASA researchers named their cluster Beowulf, after the lean, mean hero of medieval legend who defeated the giant monster Grendel by ripping off one of the creature's arms. Since then, the name has been widely adopted to refer to any low-cost cluster constructed from commercially available PCs. In 1996 two successors to the original Beowulf cluster appeared: Hyglac (built by researchers at the California Institute of Technology and the Jet Propulsion Laboratory) and Loki (constructed at Los Alamos National Laboratory). Each cluster integrated 16 Intel Pentium Pro microprocessors and showed sustained performance of over one gigaflops at a cost of less than \$50,000, thus satisfying NASA's original goal.

The Beowulf approach seemed to be the perfect computational solution to our problem of mapping the ecoregions of the U.S. A single workstation could handle the data for only a few states at most, and we couldn't assign different regions of the country to separate workstations—the en-

THE AUTHORS

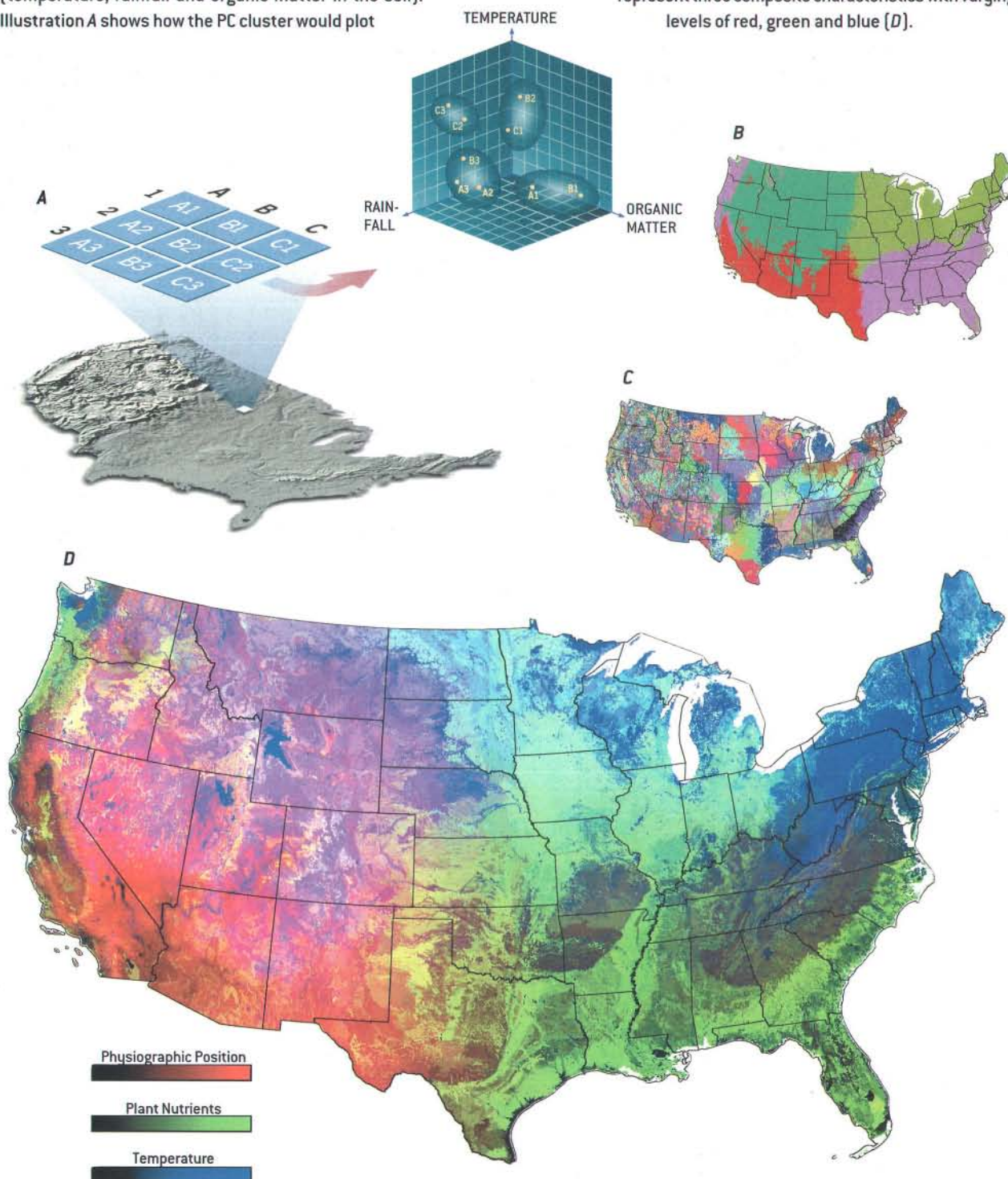
WILLIAM W. HARGROVE, FORREST M. HOFFMAN and **THOMAS STERLING** are pioneers of Beowulf computing. Hargrove, who works in the computational physics and engineering division at Oak Ridge National Laboratory in Tennessee, is a landscape ecologist with big problems and too much data. Hoffman, a computer specialist in the environmental sciences division at ORNL, spends his spare time building supercomputers in his basement. Sterling, who created the first Beowulf cluster while at the NASA Goddard Space Flight Center, is at the California Institute of Technology's Center for Advanced Computing Research and is a principal scientist at the Jet Propulsion Laboratory.

MAKING MAPS WITH THE STONE SOUPERCOMPUTER

TO DRAW A MAP of the ecoregions in the continental U.S., the Stone SouperComputer compared 25 environmental characteristics of 7.8 million one-square-kilometer cells. As a simple example, consider the classification of nine cells based on only three characteristics [temperature, rainfall and organic matter in the soil].

Illustration A shows how the PC cluster would plot

the cells in a three-dimensional data space and group them into four ecoregions. The four-region map divides the U.S. into recognizable zones (*illustration B*); a map dividing the country into 1,000 ecoregions provides far more detail (*C*). Another approach is to represent three composite characteristics with varying levels of red, green and blue (*D*).



vironmental data for every section of the country had to be compared and processed simultaneously. In other words, we needed a parallel-processing system. So in 1996 we wrote a proposal to buy 64 new PCs containing Pentium II microprocessors and construct a Beowulf-class supercomputer. Alas, this idea sounded implausible to the reviewers at ORNL, who turned down our proposal.

Undeterred, we devised an alternative plan. We knew that obsolete PCs at the U.S. Department of Energy complex at Oak Ridge were frequently replaced with newer models. The old PCs were advertised on an internal Web site and auctioned off as surplus equipment. A quick check revealed hundreds of outdated computers waiting to be discarded this way.

cation among the nodes and therefore can be solved very quickly by parallel-processing systems.

Anyone building a Beowulf cluster must make several decisions in designing the system. To connect the PCs, researchers can use either standard Ethernet networks or faster, specialized networks, such as Myrinet. Our lack of a budget dictated that we use Ethernet, which is free. We chose one PC to be the front-end node of the cluster and installed two Ethernet cards into the machine. One card was for communicating with outside users, and the other was for talking with the rest of the nodes, which would be linked in their own private network. The PCs coordinate their tasks by sending messages to one another. The two

disk, lots of memory or (best of all) an upgraded motherboard donated to us by accident? Often all we found was a tired old veteran with a fan choked with dust.

Our room at Oak Ridge turned into a morgue filled with the picked-over carcasses of dead PCs. Once we opened a machine, we recorded its contents on a "toe tag" to facilitate the extraction of its parts later on. We developed favorite and least favorite brands, models and cases and became adept at thwarting passwords left by previous owners. On average, we had to collect and process about five PCs to make one good node.

As each new node joined the cluster, we loaded the Linux operating system onto the machine. We soon figured out how to eliminate the need to install a key-

Our room at Oak Ridge **TURNED INTO A MORGUE** filled with the picked-over carcasses of dead PCs.

Perhaps we could build our Beowulf cluster from machines that we could collect and recycle free of charge. We commandeered a room at ORNL that had previously housed an ancient mainframe computer. Then we began collecting surplus PCs to create the Stone SouperComputer.

A Digital Chop Shop

THE STRATEGY BEHIND parallel computing is "divide and conquer." A parallel-processing system divides a complex problem into smaller component tasks. The tasks are then assigned to the system's nodes—for example, the PCs in a Beowulf cluster—which tackle the components simultaneously. The efficiency of parallel processing depends largely on the nature of the problem. An important consideration is how often the nodes must communicate to coordinate their work and to share intermediate results. Some problems must be divided into myriad minuscule tasks; because these fine-grained problems require frequent internode communication, they are not well suited for parallel processing. Coarse-grained problems, in contrast, can be divided into relatively large chunks. These problems do not require much communi-

cation among the nodes and therefore can be solved very quickly by parallel-processing systems.

Most popular message-passing libraries are message-passing interface (MPI) and parallel virtual machine (PVM), which are both available at no cost on the Internet. We use both systems in the Stone SouperComputer. Many Beowulf clusters are homogeneous, with all the PCs containing identical components and microprocessors. This uniformity simplifies the management and use of the cluster but is not an absolute requirement. Our Stone SouperComputer would have a mix of processor types and speeds because we intended to use whatever surplus equipment we could find. We began with PCs containing Intel 486 processors but later added only Pentium-based machines with at least 32 megabytes of RAM and 200 megabytes of hard-disk storage.

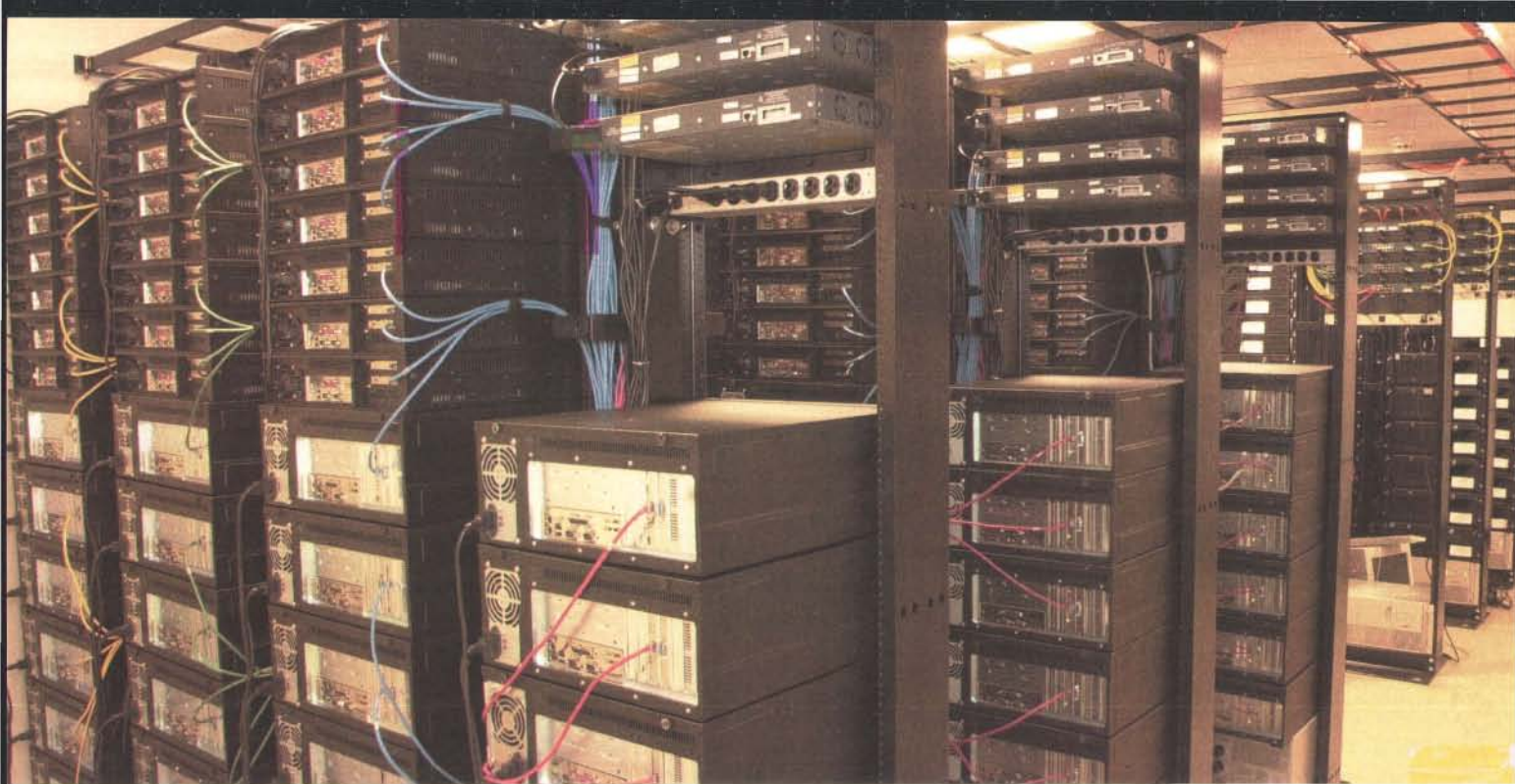
It was rare that machines met our minimum criteria on arrival; usually we had to combine the best components from several PCs. We set up the digital equivalent of an automobile thief's chop shop for converting surplus computers into nodes for our cluster. Whenever we opened a machine, we felt the same anticipation that a child feels when opening a birthday present: Would the computer have a big

board or monitor for each node. We created mobile "crash carts" that could be wheeled over and plugged into an ailing node to determine what was wrong with it. Eventually someone who wanted space in our room bought us shelves to consolidate our collection of hardware. The Stone SouperComputer ran its first code in early 1997, and by May 2001 it contained 133 nodes, including 75 PCs with Intel 486 microprocessors, 53 faster Pentium-based machines and five still faster Alpha workstations, made by Compaq.

Upgrades to the Stone SouperComputer are straightforward: we replace the slowest nodes first. Each node runs a simple speed test every hour as part of the cluster's routine housekeeping tasks. The ranking of the nodes by speed helps us to fine-tune our cluster. Unlike commercial machines, the performance of the Stone SouperComputer continually improves, because we have an endless supply of free upgrades.

Parallel Problem Solving

PARALLEL PROGRAMMING requires skill and creativity and may be more challenging than assembling the hardware of a Beowulf system. The most common



model for programming Beowulf clusters is a master-slave arrangement. In this model, one node acts as the master, directing the computations performed by one or more tiers of slave nodes. We run the same software on all the machines in the Stone SouperComputer, with separate sections of code devoted to the master and slave nodes. Each microprocessor in the cluster executes only the appropriate section. Programming errors can have dramatic effects, resulting in a digital train wreck as the crash of one node derails the others. Sorting through the wreckage to find the error can be difficult.

Another challenge is balancing the processing workload among the cluster's PCs. Because the Stone SouperComputer contains a variety of microprocessors with very different speeds, we cannot divide the workload evenly among the nodes: if we did so, the faster machines would sit idle for long periods as they waited for the slower machines to finish processing. Instead we developed a programming algorithm that allows the master node to send more data to the faster slave nodes as they complete their tasks. In this load-balancing arrangement, the faster PCs do most of the work, but the slower machines still contribute to the system's performance.

Our first step in solving the ecoregion mapping problem was to organize the

COMPUTING CLUSTER at the American Museum of Natural History in New York City contains 560 Pentium III microprocessors. Researchers use the system to study evolution and star formation.

enormous amount of data—the 25 environmental characteristics of the 7.8 million cells of the continental U.S. We created a 25-dimensional data space in which each dimension represented one of the variables (average temperature, precipitation, soil characteristics and so on). Then we identified each cell with the appropriate point in the data space [see *illustration A* on page 54]. Two points close to each other in this data space have, by definition, similar characteristics and thus are classified in the same ecoregion. Geographic proximity is not a factor in this kind of classification; for example, if two mountaintops have very similar environments, their points in the data space are very close to each other, even if the mountaintops are actually thousands of miles apart.

Once we organized the data, we had to specify the number of ecoregions that would be shown on the national map. The cluster of PCs gives each ecoregion an initial “seed position” in the data space. For each of the 7.8 million data points, the system determines the closest seed position and assigns the point to the corresponding ecoregion. Then the cluster finds the centroid for each ecoregion—the average position of all the points assigned to

the region. This centroid replaces the seed position as the defining point for the ecoregion. The cluster then repeats the procedure, reassigning the data points to ecoregions depending on their distances from the centroids. At the end of each iteration, new centroid positions are calculated for each ecoregion. The process continues until fewer than a specified number of data points change their ecoregion assignments. Then the classification is complete.

The mapping task is well suited for parallel processing because different nodes in the cluster can work independently on subsets of the 7.8 million data points. After each iteration the slave nodes send the results of their calculations to the master node, which averages the numbers from all the subsets to determine the new centroid positions for each ecoregion. The master node then sends this information back to the slave nodes for the next round of calculations. Parallel processing is also useful for selecting the best seed positions for the ecoregions at the very beginning of the procedure. We devised an algorithm that allows the nodes in the Stone SouperComputer to determine collectively the most widely dispersed data points, which are then chosen as the seed positions. If the cluster starts with well-dispersed seed

Above all, the Beowulf concept is an **EMPOWERING FORCE**.

positions, fewer iterations are needed to map the ecoregions.

The result of all our work was a series of maps of the continental U.S. showing each ecoregion in a different color [see illustrations B and C on page 54]. We produced maps showing the country divided into as few as four ecoregions and as many as 5,000. The maps with fewer ecoregions divided the country into recognizable zones—for example, the Rocky Mountain states and the desert Southwest. In contrast, the maps with thousands of ecoregions are far more complex than any previous classification of the country's environments. Because many plants and animals live in only one or two ecoregions, our maps may be useful to ecologists who study endangered species.

In our first maps the colors of the ecoregions were randomly assigned, but we later produced maps in which the colors of the ecoregions reflect the similarity of their respective environments. We statistically combined nine of the environmental variables into three composite characteristics, which we represented on the map with varying levels of red, green and blue. When the map is drawn this way, it shows gradations of color instead of sharp borders: the lush Southeast is mostly green, the cold Northeast is mainly blue, and the arid West is primarily red [see illustration D on page 54].

MORE TO EXPLORE

Cluster Computing: Linux Taken to the Extreme. F. M. Hoffman and W. W. Hargrove in *Linux Magazine*, Vol. 1, No. 1, pages 56–59; Spring 1999.

Using Multivariate Clustering to Characterize Ecoregion Borders. W. W. Hargrove and F. M. Hoffman in *Computers in Science and Engineering*, Vol. 1, No. 4, pages 18–25; July/August 1999.

How to Build a Beowulf: A Guide to the Implementation and Application of PC Clusters. Edited by T. Sterling, J. Salmon, D. J. Becker and D. F. Savarese. MIT Press, 1999.

More information about Beowulf computing can be found at the following Web sites:

stonesoup.esd.ornl.gov/
extremelinux.esd.ornl.gov/
www.beowulf.org/
www.cacr.caltech.edu/research/beowulf/beowulf-underground.org/

Moreover, the Stone SouperComputer was able to show how the ecoregions in the U.S. would shift if there were nationwide changes in environmental conditions as a result of global warming. Using two projected climate scenarios developed by other research groups, we compared the current ecoregion map with the maps predicted for the year 2099. According to these projections, by the end of this century the environment in Pittsburgh will be more like that of present-day Atlanta, and conditions in Minneapolis will resemble those in present-day St. Louis.

The Future of Clusters

THE TRADITIONAL MEASURE of supercomputer performance is benchmark speed: how fast the system runs a standard program. As scientists, however, we prefer to focus on how well the system can handle practical applications. To evaluate the Stone SouperComputer, we fed the same ecoregion mapping problem to ORNL's Intel Paragon supercomputer shortly before it was retired. At one time, this machine was the laboratory's fastest, with a peak performance of 150 gigaflops. On a per-processor basis, the run time on the Paragon was essentially the same as that on the Stone SouperComputer. We have never officially clocked our cluster (we are loath to steal computing cycles from real work), but the system has a theoretical peak performance of about 1.2 gigaflops. Ingenuity in parallel algorithm design is more important than raw speed or capacity: in this young science, David and Goliath (or Beowulf and Grendel!) still compete on a level playing field.

The Beowulf trend has accelerated since we built the Stone SouperComputer. New clusters with exotic names—Grendel, Naegling, Megalon, Brahma, Avalon, Medusa and theHive, to mention just a few—have steadily raised the performance curve by delivering higher speeds at lower costs. As of last November, 28 clusters of PCs, workstations or servers were on the list of the world's 500 fastest computers. The LosLobos cluster at the University of New Mexico has 512 Intel Pen-

tium III processors and is the 80th-fastest system in the world, with a performance of 237 gigaflops. The Cplant cluster at Sandia National Laboratories has 580 Compaq Alpha processors and is ranked 84th. The National Science Foundation and the U.S. Department of Energy are planning to build even more advanced clusters that could operate in the teraflops range (one trillion floating-point operations per second), rivaling the speed of the fastest supercomputers on the planet.

Beowulf systems are also muscling their way into the corporate world. Major computer vendors are now selling clusters to businesses with large computational needs. IBM, for instance, is building a cluster of 1,250 servers for NuTec Sciences, a biotechnology firm that plans to use the system to identify disease-causing genes. An equally important trend is the development of networks of PCs that contribute their processing power to a collective task. An example is SETI@home, a project launched by researchers at the University of California at Berkeley who are analyzing deep-space radio signals for signs of intelligent life. SETI@home sends chunks of data over the Internet to more than three million PCs, which process the radio-signal data in their idle time. Some experts in the computer industry predict that researchers will eventually be able to tap into a "computational grid" that will work like a power grid: users will be able to obtain processing power just as easily as they now get electricity.

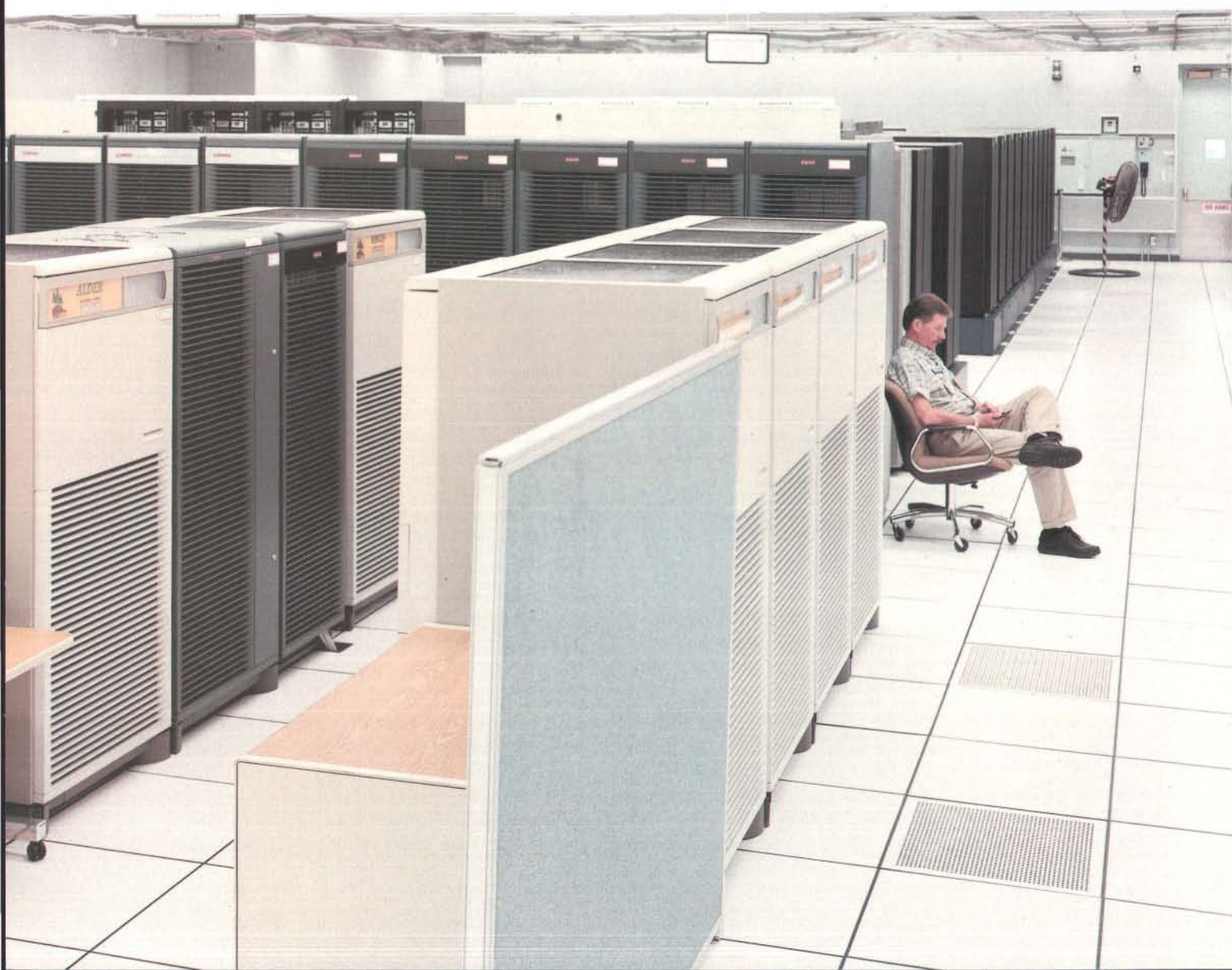
Above all, the Beowulf concept is an empowering force. It wrests high-level computing away from the privileged few and makes low-cost parallel-processing systems available to those with modest resources. Research groups, high schools, colleges or small businesses can build or buy their own Beowulf clusters, realizing the promise of a supercomputer in every basement. Should you decide to join the parallel-processing proletariat, please contact us through our Web site (<http://extremelinux.esd.ornl.gov/>) and tell us about your Beowulf-building experiences. We have found the Stone Soup to be hearty indeed.

How to Build a **HYPER** computer

BY THOMAS STERLING

The simulation and ultimate solution of
humanity's major ills and most perplexing problems require
significantly faster supercomputers

Photographs by Olivier Laude



TODAY'S FASTEST supercomputers run too slowly to do tomorrow's science. Despite the ongoing revolution in communications and information processing, many computational challenges critical to the future health, welfare, security and prosperity of humankind cannot be met by even the quickest computers. Crucial advances in pivotal fields such as climatology, medicine, bioscience, controlled fusion, national defense, nanotechnology, advanced engineering and commerce depend on the development of machines that will operate at speeds at least 1,000 times faster than today's biggest supercomputers [see "Crucial Tasks for Hypercomputers," on page 61].

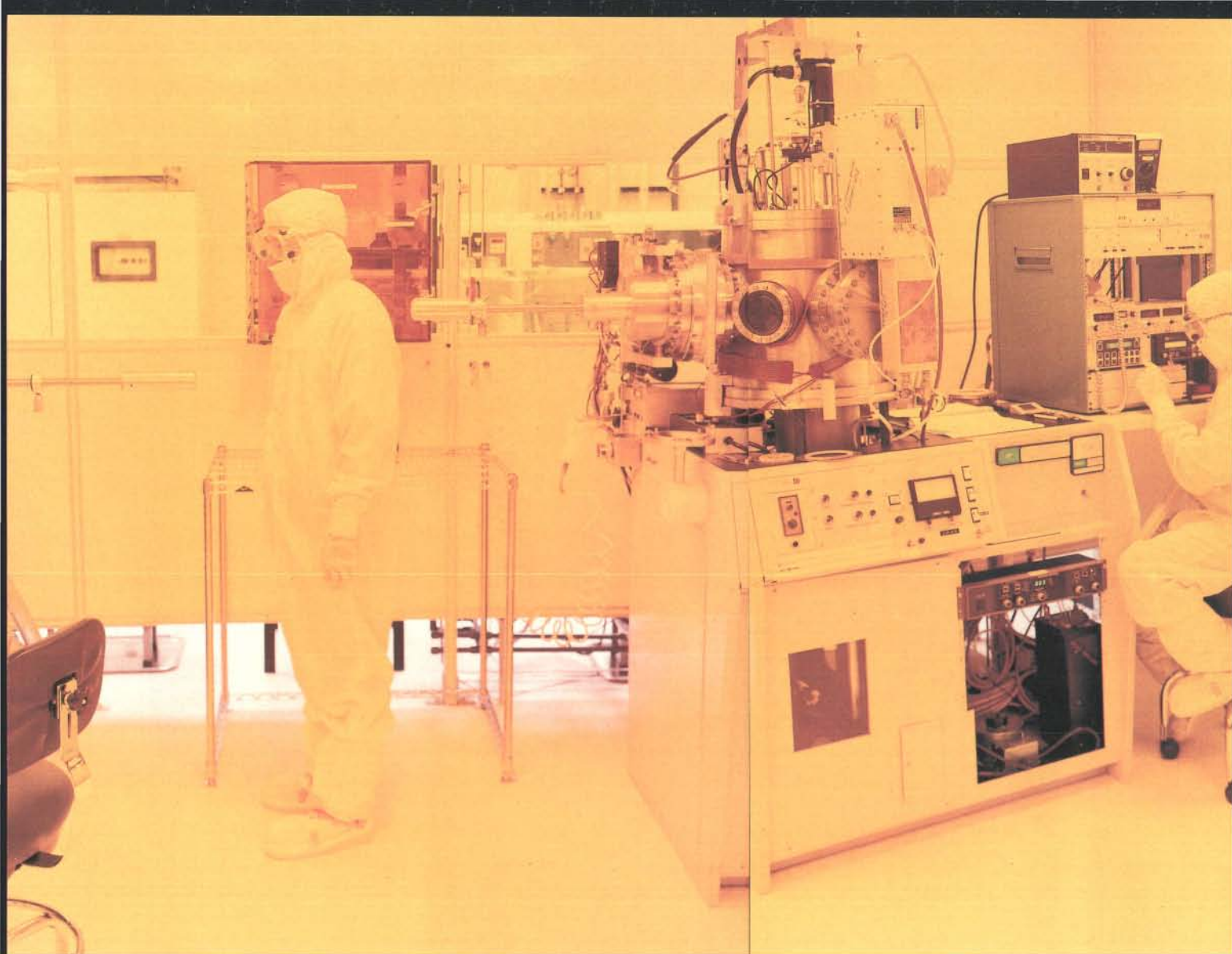
Solutions to these incredibly complex problems hinge on the ability to simulate and model their behavior with a high degree of fidelity and reliability, often over long periods. This level of performance goes far beyond that of present-day supercomputers, which at best can execute several trillion floating-point operations per second (teraflops). It could take 100 years, for example, for the largest existing system to perform a com-

plete protein-folding computation—a long-sought capability. To accomplish this kind of analysis task, researchers need hypercomputing systems that achieve at least petaflops speeds—that is, more than a quadrillion floating-point operations (arithmetical calculations) per second.

Not only do current high-end computers run too slowly, they cost too much. The three-teraflops (peak performance) ASCI (Accelerated Strategic Computer Initiative) Blue systems that are dedicated to the stewardship of the U.S. nuclear stockpile cost approximately \$120 million each. That's equivalent to a price/performance factor of \$40 per peak megaflops (million flops), which is more than 10 times greater than the price/performance of a premium personal computer. High-end computers impose indirect costs as well. Annual payments for

QUICK COMPUTER: The three-teraflops ASCI [Accelerated Strategic Computer Initiative] Blue system at Lawrence Livermore National Laboratory helps to maintain the nation's nuclear weapons stockpile.





FAST CHIPS: Engineers at TRW's Space Park facility in Redondo Beach, Calif., use sputtering machines to deposit thin superconducting films on silicon wafers as part of the fabrication of prototype superconducting processor chips.

the electrical power to operate such systems can easily exceed \$1 million. Housing their oversize footprints can also add significant expense. Paying crack programmers to write the complex code for these machines is yet another cost.

Despite their impressive processing speeds, high-end systems do not make good use of the computing resources they have, resulting in surprisingly low efficiency levels. Twenty-five percent efficiency is not uncommon, and efficiencies have dropped as low as 1 percent when addressing certain applications.

The hybrid technology multithread-

ed (HTMT) system is a new class of computer that offers 100 times the capability of present high-end machines for roughly the same cost, power usage and floor space. Further development could bring the technology beyond a quadrillion flops to trans-petaflops territory—1,000 times the performance of today's best systems or more. To achieve these goals, a multi-institutional, interdisciplinary team has created a computer architecture able to harness various advanced processing, memory and communications technologies, leveraging their strengths and complementing their limitations. The basic elements of HTMT have been developed with financial support from NASA, the National Security Agency, the National Science Foundation and the Defense Advanced Research Projects Agency; actual construction awaits further governmental funding.

Ironically, it is the very success of computing technology that reveals its limitations. Back in the late 1970s, personal computers could barely play Pong. A system capable of executing a major science problem of the day at a performance level of a few tens of megaflops could cost \$40 million or more. In contrast, PCs now priced at less than \$2,000 can outperform those machines.

Historically, the supercomputer industry has pushed the frontiers of processing performance with a combination of advanced technology and architectures customized to address specific problems. The unfortunate side effect has been high price tags. Exorbitant costs and lengthy development times have kept the market for such systems relatively flat while other segments of the computer industry have grown explosively. With these costs forcing up the price to the cus-

tomers, the overall supercomputer market and corporate investment in the technology have remained limited, producing a classic commercial death spiral.

Even when alternative approaches have been tried—including custom vector computer architectures (which efficiently perform a single operation on a list of numbers using pipelined memory access and arithmetic functional units) as well as massively parallel systems integrating large arrays of cooperating microprocessors—the costs of such systems have remained high while operational efficiencies for many applications have suffered. In the past two or three years, a number of groups have built highly parallel general-purpose computers with peak performance levels of more than a teraflops. Yet low efficiency levels mean that little

of this processing capability can be brought to bear on real-world applications. As a result, commodity clusters—networked arrays of standard computing subsystems—are perceived as the only economically viable pathway: they require little additional development in spite of the programming difficulties and communications delays inherent in using clustered systems.

Research on new classes of petaflops-capable systems has been under way since the mid-1990s. Engineers have been attacking the speed problem on all fronts, pursuing various technology paths to such machines. With sufficient R&D support, all can be accomplished within this decade [see “Five Routes to Ultrafast Processing,” on the next page]. Although each method has its strengths and weaknesses,

one of the most widely applicable is the HTMT design.

HTMT exploits a diverse array of advanced technologies within a single flexible and optimized system. The project attempts to achieve efficient trans-petaflops performance by incorporating superfast processors, high-capacity communications links, high-density memory storage and other soon-to-mature technologies in a dynamic, adaptive architecture.

No matter what course they take, designers of trans-petaflops systems all face three challenges. First, they must find a way to aggregate sufficient processing, memory and communications resources to achieve the targeted peak-computing capabilities despite practical constraints of size, cost and power. The second goal is to attain reasonable operational effi-

Crucial Tasks for Hypercomputers

Many intricate scientific problems with enormous social and political implications await solutions that can be processed only on computers that can execute more than a quadrillion floating-point operations per second—trans-petaflops performance.

Climate Modeling

Perhaps the most critical issue facing the earth's inhabitants is the need for accurate predictive scenarios for both short- and long-term weather changes. First, trans-petaflops computers could integrate the huge quantities of satellite data into detailed maps. The mapped data could then be used to simulate and model the chaotic and interrelated behaviors of the elements of our global climate system, allowing accurate predictions.

Controlled Fusion

Both an answer to the world's energy problems and a way to power spacecraft across the solar system, thermonuclear fusion's vast complexity has kept it continually just over the horizon. Trans-petaflops computers would simulate the thermal, electromagnetic and nuclear interactions of large numbers of particles in a dynamic magnetic medium to help in designing practical fusion reactors.

Medicine/Bioscience

Considerably faster computing capability could give medicine the edge in combating continuously evolving diseases. This job requires molecular-level analysis to achieve nearly instantaneous drug design, including exploring complex protein folding.

Agriculture

To feed the earth's ever growing population, rapid computation will help develop new genetically engineered crops and solve the complex problems involved in managing the world's ecology.

National Defense

With the real-world testing of nuclear weapons banned, trans-petaflops machines could model the behavior of these systems to help maintain the readiness of the strategic weapons stockpile. Real-time decryption of increasingly complicated secret codes is one key to maintaining national defense.

Commerce and Finance

Large-scale mining of the enormous data spaces containing business information and economic statistics will allow a more accurate simulation of commercial systems.

Nanotechnology

With digital electronics shrinking to the atomic scale, where quantum mechanics is important, chip designers can no longer model electronics using averaged physical parameters.

Advanced Engineering

Ultrafast processing will be needed to simulate the behavior of new materials and composites at the microscale. Future aircraft design and that of other complex engineered systems will benefit from the same type of detailed modeling capabilities.

Astronomy

To model the galaxy and its 100 billion stars properly, new, superfast computers will be required to analyze the complex interplay of the interstellar medium and heavier molecules.

ciencies in the face of standard degradation factors. These include latencies (time delays) across the system, contention for shared resources such as common memory and communications channels, overhead-related resource reductions caused by the need to manage and coordinate concurrent tasks and parallel resources, and wastage of computing resources (starvation) caused by insufficient task parallelism or inadequate load balancing. The third objective concerns finding ways to improve the usability of the system—a somewhat arbitrary measure comprising the issues of generality (general utility), programmability and availability.

Superconducting Processors

DURING THE PAST DECADE, digital logic has been dominated by CMOS (complementary metal oxide semiconductor) processors. CMOS technology has provided lower power and greater performance while system densities have increased at an exponential rate. Yet the fastest digital logic technology on earth is not CMOS. An altogether different technology using another kind of physics claims that title: superconducting logic.

Discovered at the beginning of the 20th century, superconductivity is the ability to conduct electricity with no resistance, a phenomenon that some materials

exhibit when cooled to cryogenic temperatures. In principle, a loop of superconducting wire can sustain an electric current forever. More important, superconducting devices exhibit quantum-mechanical behavior in macroscale electronic components and circuits. In the early 1960s researchers developed a nonlinear switching device based on superconductivity called the Josephson junction, which was found to have exceptional speeds.

The HTMT hypercomputer design will employ high-speed superconducting logic processors based on Josephson junction technology. In rapid single-flux quantum (RSFQ) technology, supercon-

Five Routes to Ultrafast Processing

One approach to attaining trans-petaflops computing performance (more than a quadrillion floating-point operations per second) is to use a hybrid architecture combining several soon-to-be-available advanced technologies (see accompanying article). Here are five other technical pathways to achieving that goal.

NAME	METHOD	EXAMPLE	BEST APPLICATIONS
1 SPECIAL-PURPOSE ARCHITECTURE OR SYSTOLIC ARRAY	Specially designed hardware and software that mirror the abstract problem to be solved. Runs in parallel with a fast data pipeline to speed computation	Grape Project (University of Tokyo)	Huge multibody calculations, stellar cluster simulation, bioinformatics
2 CELLULAR AUTOMATA	Finite-state machine in which many relatively simple computing cells in a large 2-D or 3-D matrix operate in lockstep during each clock cycle. Each cell's actions depend on its internal state and those of its nearest neighbors	Never fully executed	Computational fluid dynamics, diffusion simulations
3 PROCESSOR-IN-MEMORY (PIM) ARCHITECTURE	With a good deal of processing and memory on each chip, the system logic sees all the bits coming out of the dynamic random-access memory (DRAM) at the same time. There's lots of memory access and little delay in data transmission speed processing during each cycle	IRAM (University of California at Berkeley) Blue Gene (IBM)	Image processing, data encryption, rapid database searches, protein-folding modeling
4 BEOWULF OR CLUSTER ARCHITECTURES	A high-bandwidth mesh system interconnects many low-cost, commodity processors (each a partial-system-on-a-chip device) in a high-density array	GigAssembler software (International Human Genome Sequencing Consortium)	Wide range of problems; deciphering the human genome
5 DISTRIBUTED COMPUTING OR MEGACOMPUTING ARCHITECTURES	Harness the unused computing cycles on the estimated 500 million personal computers linked to the Internet. Inefficient communications is a drawback	SETI@home (Serendip Project)	Huge parallel problems such as Monte Carlo simulations and monitoring the function of the Internet

Not only do current **HIGH-END SUPERCOMPUTERS** run too slowly, they **COST TOO MUCH** and use too much power.

ducting loops store information as tiny magnetic flux quanta (by discrete current levels). The loops, called superconducting quantum interference devices, or SQUIDs, are simple mechanisms originally developed as sensing devices that comprise two Josephson junctions connected by an inductor, which is like a solenoid. With both Josephson junctions operating, a current injected into the loop will continue indefinitely. SQUIDs exhibit the interesting characteristic of having distinct states of operation: they may contain no current, sustain the basic current, or have a current that is some integral multiple times the basic current but nothing in between. This remarkable property results from quantum-mechanical effects. To represent the 0's and 1's of digital code, RSFQ logic gates use discrete currents (or fluxes) rather than distinct voltage levels. When cooled to a temperature of four kelvins, these units can operate at more than 770 gigahertz, the fastest (single-gate) processing speeds ever achieved and approximately 100 times quicker than conventional CMOS logic.

RSFQ technology will allow the hybrid computing system to run nominally at from 100 to 200 gigaflops (billion flops) per processor as opposed to a few gigaflops, as in standard CMOS processors. In addition, the minuscule and packetized nature of magnetic flux quanta in RSFQ devices cuts crosstalk and power consumption by a couple of orders of magnitude. This rapidly maturing technology reduces parallelism requirements, cost, power demand and system size.

Boosting Efficiency

WITH SUPERFAST processors in place, HTMT seeks to make efficient use of their powerful capabilities. Those processors should spend their time doing little else but computations. Conventional approaches such as commodity clusters require large-scale tasks to be run on similarly large-scale computational nodes. Often a computational node on a conventional system must wait while a remote request to an-

other node is being serviced. Unless operators exactly balance the workload, some nodes will continue to compute while others, having finished their jobs, will stall. Even when engineers employ load-balancing software techniques, the overhead required for accomplishing this function can reduce efficiency.

Unlike any other computer architecture, HTMT revolutionizes the relation between the processing system and the memory system. In ordinary multiprocessor systems, the computational processors manage and manipulate the "dumb" memory system; in contrast, HTMT's "smart" memory system administers the processors. HTMT and other tightly coupled parallel computers consider the workload on the processing elements and make on-the-fly decisions as to which part of a task should be performed by what hardware. In doing so, the processors work out of their local registers and some high-speed buffer memories, thus avoiding having to reach too far out into the system. The result is a drop in latency problems. The processors do not spend time managing memory resources, which are just wasted processing cycles that add to overhead; these logistical decisions are made by the small low-cost processors in the memory.

The HTMT design attacks the problem of latency in two ways. First, the system employs a dynamic, adaptive resource management scheme based on a multithreaded architecture that enables HTMT to switch from one stream of instructions to another within a single cycle. Whereas most computers operate with one stream of instructions, HTMT will feature multiple instruction streams. By using overlapping communications, the processors can work on many out-

standing requests simultaneously. Say a superconducting processor needs to load information from a cache or a high-speed buffer, a procedure that will take many 10-picosecond cycles. As this request is served by the memory system, the processor can switch to another data stream to find operations that can be performed immediately.

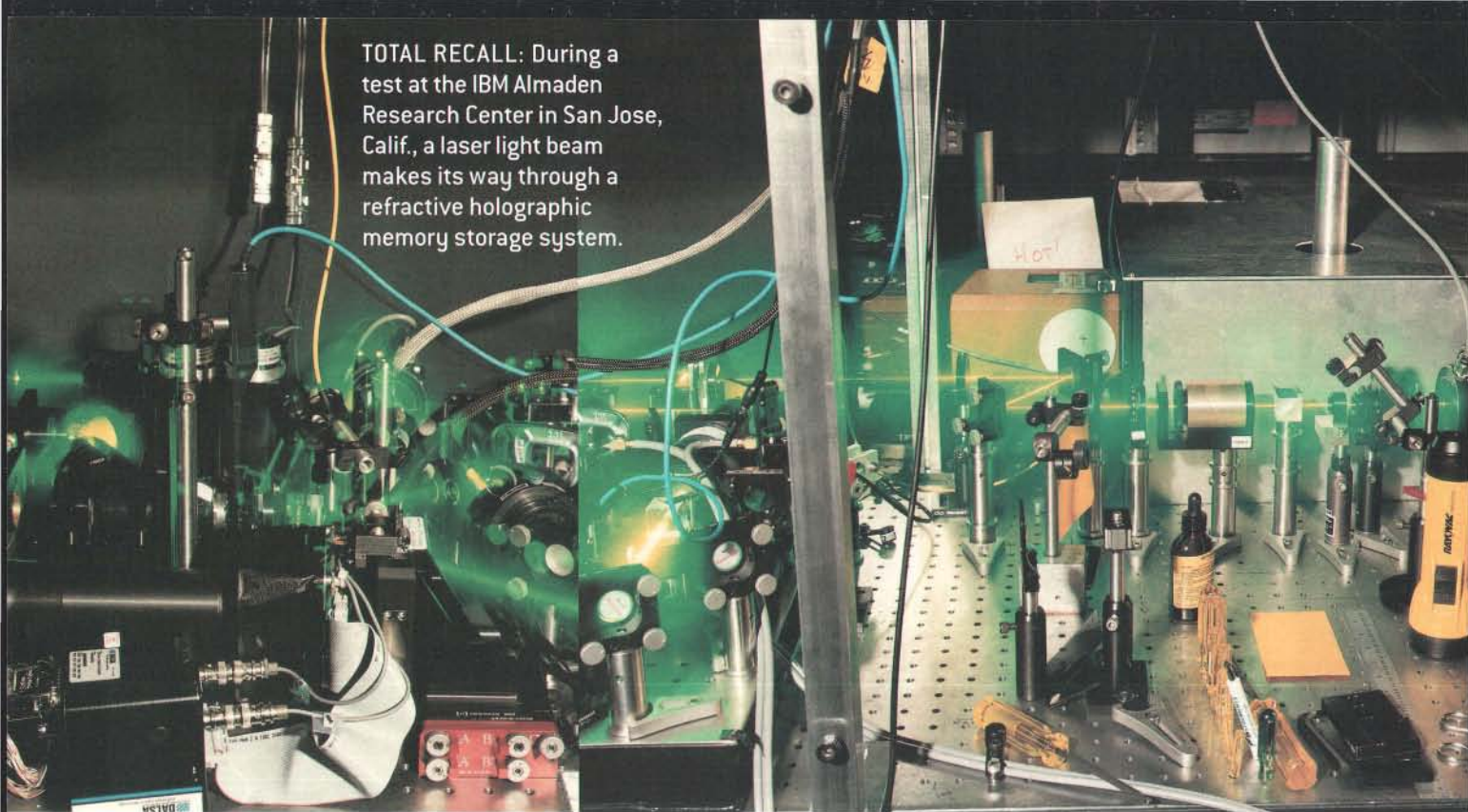
The second way HTMT will handle the latency issue is by employing processor-in-memory technology (PIM), wherein small secondary satellite or taxi logic processors are placed in its memory devices. A few years ago fabrication advances allowed CMOS logic and dynamic random-access memory (DRAM) cells to be put onto the same silicon die, permitting them to be closely integrated. These cheap devices deal with overhead—that is, they manipulate information in the memory, again allowing the superconducting processors to focus on computing. PIM processing technology can also handle memory-intensive functions such as data gathers—collecting needed information from various locations and placing it in one dense object—as well as carrying out the reverse operation of data scatters—distributing the information to the correct locations.

Although the technologies and architecture incorporated by the HTMT system may be innovative, the means of managing these resources and the computing discipline employed by HTMT are truly revolutionary. The system will use new percolation techniques in which the PIM processors decide when a new piece of work should be performed. They will determine when to migrate all information that needs to be executed up to rapid-access buffer memories near the high-speed superconducting processors.

THE AUTHOR

THOMAS STERLING holds a joint appointment at the NASA Jet Propulsion Laboratory's High Performance Computing group, where he is a principal scientist, and the California Institute of Technology's Center for Advanced Computing Research, where he is a faculty associate. For the past 20 years, Sterling has carried out research on parallel-processing hardware and software systems for high-performance computing. Since 1994 he has been a leader in the national petaflops initiative. He heads the hybrid technology multithreaded architecture research project.

TOTAL RECALL: During a test at the IBM Almaden Research Center in San Jose, Calif., a laser light beam makes its way through a refractive holographic memory storage system.



For example, when a specific subroutine is required, it and the special information it needs to execute its function will be moved up to the processors. This proactive method of prestaging necessary information is a way to avoid creating long latency delays in connecting to the main memory. The technique also frees the high-speed processors from having to

perform logistical overhead operations, because they are not needed to bring information to the processing sites.

Improving Usability

THE THIRD MAJOR challenge to transpetaflops computing concerns the usability of the system: researchers must increase its generality (to ensure that it can

handle a wide variety of problems), make it easier to program, and boost its availability, or uptime. HTMT addresses these issues in several ways.

By using a global name space in a shared-memory computing structure, every processor can “see” all of the memory. This method is more general than typical distributed- (or fragmented-) memory computing techniques because it provides efficient access by any processor to all data without having to engage software routines on a remote processor to assist in the data transfer. More actions can be performed simultaneously, speeding execution. In addition, by letting the system conduct dynamic rescheduling—responding to run-time information—it can perform certain computations more effectively, a capability that adds to its generality. And because this arrangement is closer to the way computational scientists think about their problems, programming the system is more intuitive. Typically programmers must determine beforehand how a problem should be handled by a system, a complex and laborious task. But an HTMT system makes many of these decisions by itself, thereby helping to alleviate one of the biggest difficulties in working with large computers—programming them.

The hybrid computing system will

Key Concepts/*Hypercomputing*

Contention—Time delay created when two processors try to access a shared resource simultaneously

Latency—Delay caused by the time it takes for a remote request to be serviced or for a message to travel between two processing nodes

Load Balancing—Distributing work evenly so that all processing nodes are kept occupied as the program is executed

Overhead—Time spent on noncomputational functions such as the logistical management of parallel resources and concurrent tasks

Percolation—Method of managing tasks and data movement without incurring delays caused by overhead, latency, contention or starvation

Processor-in-Memory (PIM)—Integrated circuits that contain both memory and logic on the same chip

Starvation—Wastage of computing resources caused by insufficient program parallelism or poor load balancing

Wave Division Multiplexing (WDM)—Method by which the effective bandwidth of an optical channel can be increased by using optical signals with different wavelengths

The hybrid computer **REVOLUTIONIZES** the relation between the **PROCESSING** system and the **MEMORY** system.

provide greater availability to users through the use of higher-capability sub-components, allowing it to achieve the same level of performance with fewer parts. This parts reduction increases the mean time between failures of the entire system, thus boosting operational uptime.

Holographic Memory Storage

ANOTHER INNOVATIVE aspect of the HTMT system will be its use of high-density-capacity holographic memory storage devices. This alternative to the semiconductor-based DRAM is being explored by academic and industrial research laboratories and should provide superior storage density as well as lower power consumption and costs.

Holographic storage systems use light-sensitive materials to accumulate large blocks of data. Photorefractive and spectral hole-burning techniques represent two distinct approaches. In photorefractive storage, a plane of data modulates a laser beam (signal) that interferes with a reference beam in a small rectangular block of a storage material such as lithium niobate. The hologram results from the electro-optic effect that occurs when local electric fields are created by trapped, spatially distributed charge carriers excited by the interfering beams. Many data blocks may be stored in the same target material. They are differentiated by varying either the angle of incidence or the wavelength of the laser beam. The spectral hole-burning technique relies on a non-linear response of a storage material to optical stimuli. Data are represented by changes in the photosensitive medium's absorption spectrum. Many bits can be stored at a given spatial location.

Photorefractive methods are more far advanced. But in the long-term, spectral hole-burning technology may yield significantly higher memory density. Typical holographic devices currently feature access times of several milliseconds—approximately the same as conventional secondary storage devices such as hard disks and CD-ROM drives. But advanced

techniques employing tunable lasers or arrays of laser diodes each set at a slightly different angle to one another are expected to yield access times of a few tens of microseconds. Although these access times are about two orders of magnitude longer than that of DRAM, their data bandwidths are the same or greater, and the systems are about 100 times faster than conventional disk drives. Storage capacities of 10 gigabits or more in blocks as small as a few cubic centimeters are expected within the next decade.

Optical Communications

TO CONNECT THE SPEEDY superconducting processors and high-density holographic memory systems in a network, HTMT will use high-capacity optical data pipelines. Instead of employing electrons in metal wires, HTMT will speed communications by using photons in fiber-optic cables. Wires can easily handle hundreds of megabits per second, and speeds of a few gigabits per second (gbps) can be achieved by using differential pairs of input/output pins (one goes up while the other goes down). But it could take tens of millions of wires to supply all the global communications bandwidth required of systems operating in the petaflops regime. With modulated lasers, digital light signals can transmit at up to 10 gbps per channel or more in conventional optical communications systems.

Employing multiple wavelengths (or colors) of light carrying digital information dramatically improves fiber-optic bandwidth or channel capacity. HTMT will use an advanced optical transmission

system called wave division multiplex (WDM) communications. It should provide about 100 times the per-channel bandwidth of the best conventional metal-wire communications systems. WDM allows separate digital signals, each with its own dedicated light wavelength, to travel together through the same channel. The number of different wavelengths that can be simultaneously transmitted through a single channel has grown to around 100 in recent years, and in time this figure may rise further. With improved receiver, transmitter and switch technology now in development, switching rates of 50 megahertz or more will soon be possible. Still experimental devices may bring about rates on the order of one gigahertz in the future. This capacity level would be sufficient to manage the huge information flow of a petaflops-scale computing system.

These next-generation hypercomputers would offer an important tool for exploring the world's most pressing problems, including global warming, disease epidemics and cleaner energy. In 1999 the President's Information Technology Advisory Committee strongly recommended financial support for these kinds of projects. Research groups have demonstrated that HTMT technologies could be the best route to trans-petaflops performance. Proper funding is all that is needed to put these systems in place. ■

This article is the first in a two-part series on next-generation supercomputers. The second part, "The Do-It-Yourself Supercomputer," will appear in the August issue.

MORE TO EXPLORE

Challenges of Future High-End Computing. David H. Bailey in *High Performance Computer Systems and Applications*. Edited by Jonathan Schaeffer. Kluwer Academic Publishers, 1998. Preprint available at www.nersc.gov/~dhbailey/dhbpapers/future.pdf

In Pursuit of a Quadrillion Operations per Second. Thomas Sterling in *NASA HPCC Insights*, No. 5; April 1998. Available at www.hpcc.nasa.gov/insights/vol5/petaflop.htm

The author's Web site: www.cacr.caltech.edu/~tron/

A Hybrid Technology Multithreaded (HTMT) Computer Architecture for Petaflops Computing. Thomas Sterling. On the JPL/NASA Project HTMT Web site at <http://htmt.jpl.nasa.gov/intro.html>

NASA high-performance computing and communications Web site: www.hq.nasa.gov/hpcc/petaflops/



VIRTUALLY THERE

*Three-dimensional tele-immersion may
eventually bring the world to your desk*

BY JARON LANIER

PHOTOGRAPH BY DAN WINTERS



JARON LANIER, physically located in Armonk, N.Y., as he appears on a tele-immersion screen in Chapel Hill, N.C.

*Like many researchers, I am a frequent but reluctant user of videoconferencing. **Human interaction** has both verbal and nonverbal elements, and videoconferencing seems precisely configured to confound the nonverbal ones. It is impossible to make eye contact*

properly, for instance, in today's videoconferencing systems, because the camera and the display screen cannot be in the same spot. This usually leads to a deadened and formal affect in interactions, eye contact being a nearly ubiquitous subconscious method of affirming trust. Furthermore, participants aren't able to establish a sense of position relative to one another and therefore have no clear way to direct attention, approval or disapproval.

Tele-immersion, a new medium for human interaction enabled by digital technologies, approximates the illusion that a user is in the same physical space as other people, even though the other participants might in fact be hundreds or thousands of miles away. It combines the display and interaction techniques of virtual reality with new vision technologies that transcend the traditional limitations of a camera. Rather than merely observing people and their immediate environment from one vantage point, tele-immersion stations convey them as "moving sculptures," without favoring a single

point of view. The result is that all the participants, however distant, can share and explore a life-size space.

Beyond improving on videoconferencing, tele-immersion was conceived as an ideal application for driving network-engineering research, specifically for Internet2, the primary research consortium for advanced network studies in the U.S. If a computer network can support tele-immersion, it can probably support any other application. This is because tele-immersion demands as little delay as possible from flows of information (and as little inconsistency in delay), in addition to the more common demands for very large and reliable flows.

Virtual Reality and Networks

BECAUSE TELE-IMMERSION sits at the crossroads of research in virtual reality and networking, as well as computer vision and user-interface research, a little background in these various fields of research is in order.

In 1965 Ivan Sutherland, who is widely regarded as the father of computer graph-

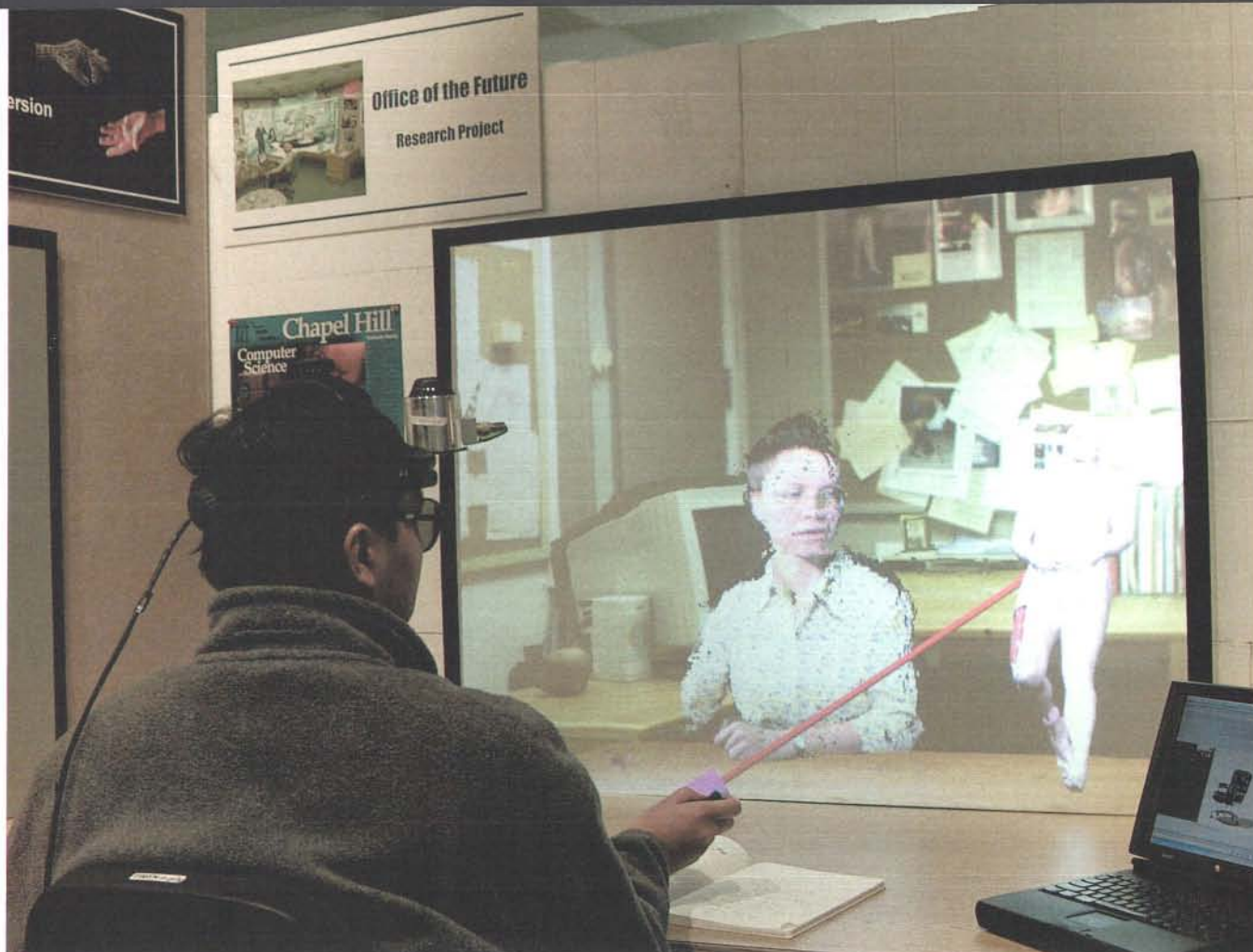
ics, proposed what he called the "Ultimate Display." This display would allow the user to experience an entirely computer-rendered space as if it were real. Sutherland termed such a space a "Virtual World," invoking a term from the philosophy of aesthetics, particularly the writings of Suzanne K. Langer. In 1968 Sutherland realized a virtual world for the first time by means of a device called a head-mounted display. This was a helmet with a pair of display screens positioned in front of the eyes to give the wearer a sense of immersion in a stereoscopic, three-dimensional space. When the user moved his or her head, a computer would quickly recompute the images in front of each eye to maintain the illusion that the computer-rendered world remained stationary as the user explored it.

In the course of the 1980s I unintentionally ended up at the helm of the first company to sell general-purpose tools for making and experiencing virtual worlds—in large part because of this magazine. *Scientific American* devoted its September 1984 issue to emerging digital technologies and chose to use one of my visual-programming experiments as an illustration for the cover.

At one point I received a somewhat panicked phone call from an editor who noticed that there was no affiliation listed for me. I explained that at the time I had no affiliation and neither did the work being described. "Sir," he informed me, "at *Scientific American* we have a strict rule that states that an affiliation must be indicated after a contributor's name." I blurted out "VPL Research" (for Visual Programming Language, or

Overview / **Tele-immersion**

- This new telecommunications medium, which combines aspects of virtual reality with videoconferencing, aims to allow people separated by great distances to interact naturally, as though they were in the same room.
- Tele-immersion is being developed as a prototype application for the new Internet2 research consortium. It involves monumental improvements in a host of computing and communications technologies, developments that could eventually lead to a variety of spin-off inventions.
- The author suggests that within 10 years, tele-immersion could substitute for many types of business travel.



Virtual Programming Language), and thus was born VPL. After the issue's publication, investors came calling, and a company came to exist in reality. In the mid-1980s VPL began selling virtual-world tools and was well known for its introduction of glove devices, which were featured on another *Scientific American* cover, in October 1987.

VPL performed the first experiments in what I decided to call "virtual reality" in the mid- to late 1980s. Virtual reality combines the idea of virtual worlds with networking, placing multiple participants in a virtual space using head-mounted displays. In 1989 VPL introduced a product called RB2, for "Reality Built for Two," that allowed two participants to share a virtual world. One intriguing implication of virtual reality is that partici-

pants must be able to see representations of one another, often known as avatars. Although the computer power of the day limited our early avatars to extremely simple, cartoonish computer graphics that only roughly approximated the faces of users, they nonetheless transmitted the motions of their hosts faithfully and thereby conveyed a sense of presence, emotion and locus of interest.

At first our virtual worlds were shared across only short physical distances, but we also performed some experiments with long-distance applications. We were able to set up virtual-reality sessions with participants in Japan and California and in Germany and California. These demonstrations did not strain the network, because only the participants' motions needed to be sent, not the entire surface of each

TELE-COLLABORATORS hundreds of miles apart consider a computer-generated medical model, which both of them can manipulate as though it were a real object. The headpiece helps the computers locate the position and orientation of the user's head; such positioning is essential for presenting the right view of a scene. In the future, the headpiece should be unnecessary.

person, as is the case with tele-immersion.

Computer-networking research started in the same era as research into virtual worlds. The original network, the Arpanet, was conceived in the late 1960s. Other networks were inspired by it, and in the 1980s all of them merged into the Internet. As the Internet grew, various "backbones" were built. A backbone is a network within a network that lets information travel over exceptionally powerful, widely shared connections to go long distances more quickly. Some notable backbones designed to support research were the NSFnet in the late 1980s and the vBNS in the mid-1990s. Each of these played a part in inspiring new applications for the Internet, such as the

JARON LANIER is a computer scientist often described as "the father of virtual reality." In addition to that field, his primary areas of study have been visual programming, simulation, and high-performance networking applications. He is chief scientist of Advanced Network and Services, a nonprofit concern in Armonk, N.Y., that funds and houses the engineering office of Internet2. Music is another of Lanier's great interests: he writes for orchestra and other ensembles and plays an extensive, exotic assortment of musical instruments—most notably, wind and string instruments of Asia. He is also well known as an essayist on public affairs.



THREE USERS in different cities can share a virtual space thanks to this telecubicle.

ased toward any particular viewpoint (a camera, in contrast, is locked into portraying a scene from its own position). Each place, and the people and things in it, has to be sensed from all directions at once and conveyed as if it were an animated three-dimensional sculpture. Each remote site receives information describing the whole moving sculpture and renders viewpoints as needed locally. The scanning process has to be accomplished fast enough to take place in real time—at most within a small fraction of a second. The sculpture representing a person can then be updated quickly enough to achieve the illusion of continuous motion. This illusion starts to appear at about 12.5 frames per second (fps) but becomes robust at about 25 fps and better still at faster rates.

Measuring the moving three-dimensional contours of the inhabitants of a room and its other contents can be accomplished in a variety of ways. As ear-

World Wide Web. Another backbone-research project, called Abilene, began in 1998, and it was to serve a university consortium called Internet2.

Abilene now reaches more than 170 American research universities. If the only goal of Internet2 were to offer a high level of bandwidth (that is, a large number of bits per second), then the mere existence of Abilene and related resources would be sufficient. But Internet2 research

tion called Advanced Network and Services, which housed and administered the engineering office for Internet2. He used the term “tele-immersion” to conjure an ideal “driver” application and asked me to take the assignment as lead scientist for a National Tele-Immersion Initiative to create it. I was delighted, as this was the logical extension of my previous work in shared virtual worlds.

Although many components, such

Seen through polarizing glasses, two walls of the cubicle dissolved into windows, revealing offices with people who **WERE LOOKING BACK AT ME.**

targeted additional goals, among them the development of new protocols for handling applications that demand very high bandwidth and very low, controlled latencies (delays imposed by processing signals en route).

Internet2 had a peculiar problem: no existing applications required the anticipated level of performance. Computer science has traditionally been driven by an educated guess that there will always be good uses for faster and more capacious digital tools, even if we don't always know in advance what those uses will be. In the case of advanced networking research, however, this faith wasn't enough. The new ideas would have to be tested on something.

Allan H. Weis, who had played a central role in building the NSFnet, was in charge of a nonprofit research organiza-

tion called Advanced Network and Services, which housed and administered the engineering office for Internet2. He used the term “tele-immersion” to conjure an ideal “driver” application and asked me to take the assignment as lead scientist for a National Tele-Immersion Initiative to create it. I was delighted, as this was the logical extension of my previous work in shared virtual worlds.

Beyond the Camera as We Know It

THE KEY IS THAT in tele-immersion, each participant must have a personal viewpoint of remote scenes—in fact, two of them, because each eye must see from its own perspective to preserve a sense of depth. Furthermore, participants should be free to move about, so each person's perspective will be in constant motion.

Tele-immersion demands that each scene be sensed in a manner that is not bi-

ly as 1993, Henry Fuchs of the University of North Carolina at Chapel Hill had proposed one method, known as the “sea of cameras” approach, in which the viewpoints of many cameras are compared. In typical scenes in a human environment, there will tend to be visual features, such as a fold in a sweater, that are visible to more than one camera. By comparing the angle at which these features are seen by different cameras, algorithms can piece together a three-dimensional model of the scene.

This technique had been explored in non-real-time configurations, notably in Takeo Kanade's work, which later culminated in the “Virtualized Reality” demonstration at Carnegie Mellon University, reported in 1995. That setup consisted of 51 inward-looking cameras mounted on a geodesic dome. Because it

was not a real-time device, it could not be used for tele-immersion. Instead videotape recorders captured events in the dome for later processing.

Ruzena Bajcsy, head of the GRASP (General Robotics, Automation, Sensing and Perception) Laboratory at the University of Pennsylvania, was intrigued by the idea of real-time seas of cameras. Starting in 1994, she worked with colleagues at Chapel Hill and Carnegie Mellon on small-scale "puddles" of two or three cameras to gather real-world data for virtual-reality applications.

Bajcsy and her colleague Kostas Daniilidis took on the assignment of creating the first real-time sea of cameras—one that was, moreover, scalable and modular so that it could be adapted to a variety of rooms and uses. They worked closely with the Chapel Hill team, which was responsible for taking the "animated sculpture" data and using computer graphics techniques to turn it into a realistic scene for each user.

But a sea of cameras in itself isn't a complete solution. Suppose a sea of cameras is looking at a clean white wall. Because there are no surface features, the cameras have no information with which to build a sculptural model. A person can look at a white wall without being confused. Humans don't worry that a wall might actually be a passage to an infinitely deep white chasm, because we don't rely on geometric cues alone—we also have a model of a room in our minds that can rein in errant mental interpretations. Unfortunately, to today's digital cameras, a person's forehead or T-shirt can present the same challenge as a white wall, and today's software isn't smart enough to undo the confusion that results.

Researchers at Chapel Hill came up with a novel method that has shown promise for overcoming this obstacle, called "imperceptible structured light," or ISL. Conventional lightbulbs flicker 50 or 60 times a second, fast enough for the flickering to be generally invisible to the human eye. Similarly, ISL appears to the human eye as a continuous source of white light, like an ordinary lightbulb, but in fact it is filled with quickly changing patterns visible only to specialized, care-

fully synchronized cameras. These patterns fill in voids such as white walls with imposed features that allow a sea of cameras to complete the measurements.

The Eureka Moment

WE WERE ABLE TO demonstrate tele-immersion for the first time on May 9, 2000, virtually bringing together three locations. About a dozen dignitaries were physically at the telecubicle in Chapel Hill. There we and they took turns sitting down in the simulated office of tomorrow. As fascinating as the three years of research leading up to this demonstration had been for me, the delight of experiencing tele-immersion was unanticipated and incomparable. Seen through a pair of polarizing glasses, two walls of the cubicle dissolved into windows, revealing other offices with other people who were looking back at me. (The glasses helped to direct a slightly different view of the scenes to each eye, creating the stereo vision effect.) Through one wall I greeted Amela Sadagic, a researcher at my lab in Armonk, N.Y. Through the other wall was Jane Mulligan, a postdoctoral fellow at the University of Pennsylvania.

Unlike the cartoonish virtual worlds I had worked with for many years, the remote people and places I was seeing were clearly derived from reality. They were not perfect by any means. There was "noise" in the system that looked something like confetti being dropped in the other people's cubicles. The frame rate was low (2 to 3 fps), there was as much as one second of delay, and only one side of the conversation had access to a tele-immersive display. Nevertheless, here was a virtual world that was not a simplistic artistic representation of the real world but rather an authentic measurement-based rendition of it.

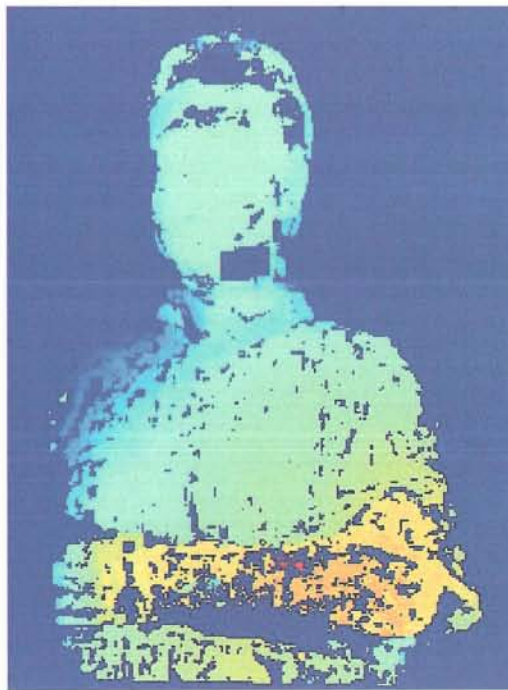
In a later demo (in October 2000) most of the confetti was

gone and the overall quality and speed of the system had increased, but the most important improvement came from researchers at Brown University led by Andries van Dam. They arrived in a tele-immersive session bearing virtual objects not derived from the physical scene. I sat across the table from Robert C. Zeleznik of Brown, who was physically at my lab in Armonk. He presented a simulated miniature office interior (about two feet wide) resting on the desk between us, and we used simulated laser pointers and other devices to modify walls and furniture in it collaboratively while we talked. This was a remarkable blending of the experience of using simulations associated with virtual reality and simply being with another person.

When Can I Use It?

BEYOND THE SCENE-CAPTURE system, the principal components of a tele-immersion setup are the computers, the network services, and the display and interaction devices. Each of these components has been advanced in the cause of tele-immersion and must advance further. Tele-immersion is a voracious consumer of computer resources. We've chosen to work with "commodity" computer components (those that are also used in common home and office products) wherever

COMPARISON OF TWO VIEWS of a person taken by the tele-immersion cameras yields this image. The colors represent the first rough calculation of the depth of the person's features.



HOW TELE-IMMERSION WORKS

In this highly simplified scheme for how a future tele-immersion scheme might work, two partners separated by 1,000 miles collaborate on a new engine design

"SEA OF CAMERAS"

Hidden cameras provide many points of view that are compared to create a three-dimensional model of users and their surroundings. The cameras can be hidden behind tiny perforations in the screen, as shown here, or can be placed on the ceiling, in which case the display screen must also serve as a selectively reflective surface.

SHARED SIMULATION OBJECTS

Simulated objects appear in the space between users. These can be manipulated as if they were working models. One stream of research in the National Tele-immersion Initiative concerns finding better techniques to combine models developed by people on opposite ends of a dialogue using incompatible local software design tools.

FOLLOWING THE FLOW OF INFORMATION

Tele-immersion depends on intense data processing at each end of a connection, mediated by a high-performance network.

FROM THE SENDER ...

Parallel processors accept visual input from the cameras and reinterpret the scene as a three-dimensional computer model.



INTERNET2



IMPERCEPTIBLE STRUCTURED LIGHT

It looks like standard white illumination to the naked eye, but it projects unnoticeably brief flickerings of patterns that help the computers make sense of otherwise featureless visual expanses.



VIRTUAL MIRROR

Users might be able check on how they and their environment appear to others through interface design features such as a virtual mirror. In this whimsical example, the male user has chosen to appear in more formal clothing than he is wearing in reality. Software to achieve this transformation does not yet exist, but early examples of related visual filtering have already appeared.

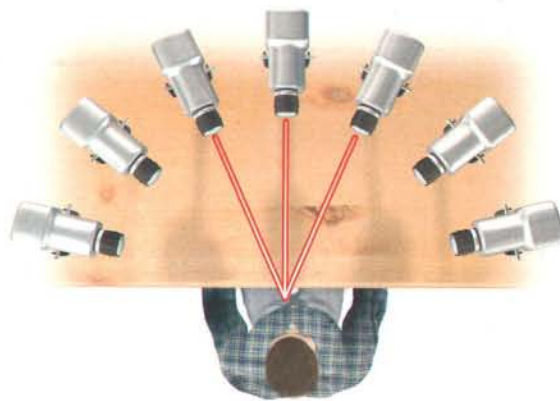
SCREEN

Current prototypes use two overlapping projections of polarized images and require users to wear polarized glasses so that each image is seen by only one eye. This technique will be replaced in the future by "autostereoscopic" displays that channel images to each eye differentially without the need for glasses.

... TO THE RECEIVER

Specific renderings of remote people and places are synthesized from the model as it is received to match the points of view of each eye of a user. The whole process repeats many times a second to keep up with the user's head motion.

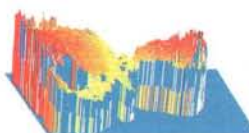
GENERATING THE 3-D IMAGE



1 An array of cameras views people and their surroundings from different angles. Each camera generates an image from its point of view many times in a second.



2 Each set of the images taken at a given instant is sorted into subsets of overlapping trios of images.

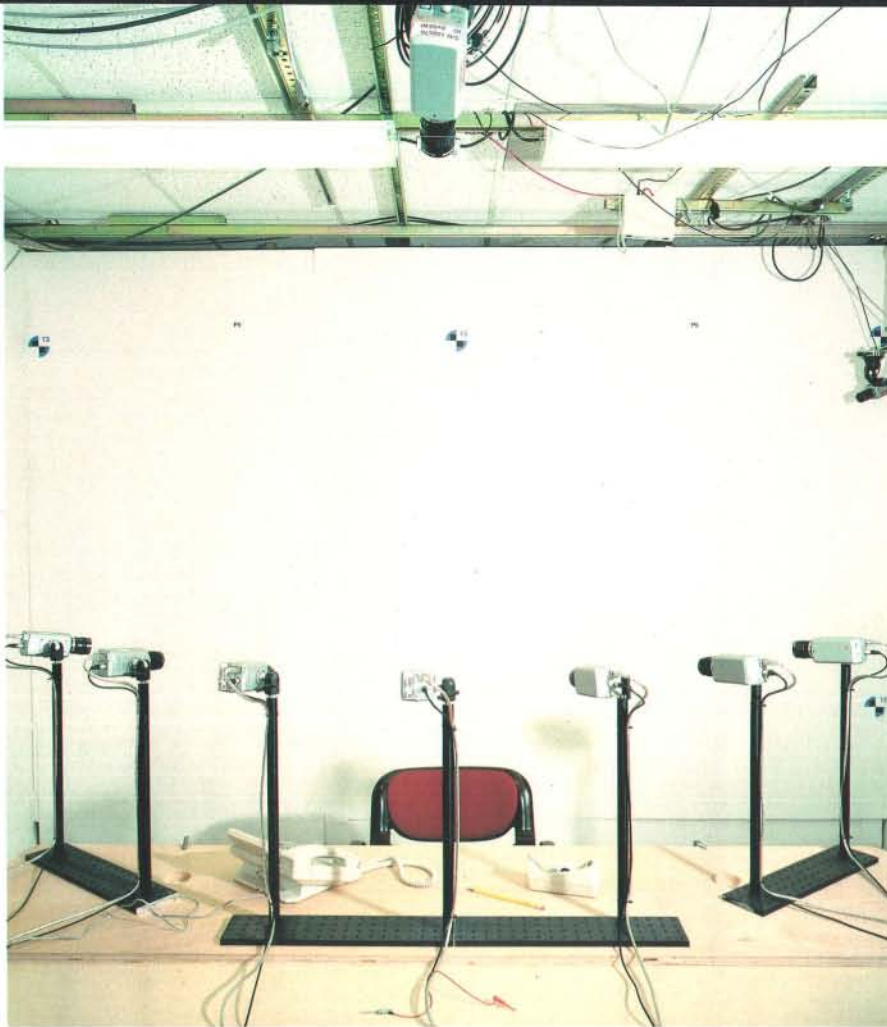


3 From each trio of images, a "disparity map" is calculated, reflecting the degree of variation among the images at all points in the visual field. The disparities are then analyzed to yield depths that would account for the differences between what each camera sees. These depth values are combined into a "bas relief" depth map of the scene.

4 All the depth maps are combined into a single viewpoint-independent sculptural model of the scene at a given moment. The process of combining the depth maps provides opportunities for removing spurious points and noise.



ILLUSTRATION BY JOE ZEFF



SEVEN CAMERAS scrutinize the user in the tele-immersion setup in Chapel Hill.

possible to hasten the day when tele-immersion will be reproducible outside the lab. Literally dozens of such processors are currently needed at each site to keep up with the demands of tele-immersion. These accumulate either as personal computers in plastic cases lined up on shelves or as circuit boards in refrigerator-size racks. I sometimes joke about the number of "refrigerators" required to achieve a given level of quality in tele-immersion.

Most of the processors are assigned to scene acquisition. A sea of cameras consists of overlapping trios of cameras. At the moment we typically use an array of

seven cameras for one person seated at a desk, which in practice act as five trios. Roughly speaking, a cluster of eight two-gigahertz Pentium processors with shared memory should be able to process a trio within a sea of cameras in approximately real time. Such processor clusters should be available later this year. Although we expect computer prices to continue to fall as they have for the past few decades, it will still be a bit of a wait before tele-immersion becomes inexpensive enough for widespread use. The cost of an eight-processor cluster is anticipated to be in the \$30,000 to \$50,000 range at introduction, and a number of those would be required for each site (one for each trio of cameras)—and this does not even account for

the processing needed for other tasks. We don't yet know how many cameras will be required for a given use of tele-immersion, but currently a good guess is that seven is the minimum adequate for casual conversation, whereas 60 cameras might be needed for the most demanding applications, such as long-distance surgical demonstration, consultation and training.

Our computational needs go beyond processing the image streams from the sea of cameras. Still more processors are required to resynthesize and render the scene from shifting perspectives as a participant's head moves during a session. Initially we used a large custom graphics computer, but more recently we have been able instead to draft commodity processors with low-cost graphics cards, using one processor per eye. Additional processors are required for other tasks, such as combining the results from each of the camera trios, running the imperceptible structured light, measuring the head motion of the user, maintaining the user interface, and running virtual-object simulations.

Furthermore, because minimizing apparent latency is at the heart of tele-immersion engineering, significant processing resources will eventually need to be applied to predictive algorithms. Information traveling through an optical fiber reaches a destination at about two thirds the speed of light in free space because it is traveling through the fiber medium instead of a vacuum and because it does not travel a straight path but rather bounces around in the fiber channel. It therefore takes anywhere from 25 to 50 milliseconds for fiber-bound bits of information to cross the continental U.S., without any allowances for other inescapable delays, such as the activities of various network signal routers.

By cruel coincidence, some critical aspects of a virtual world's responsiveness should not be subject to more than 30 to 50 milliseconds of delay. Longer delays result in user fatigue and disorientation, a degradation of the illusion and, in the worst case, nausea. Even if we had infinitely fast computers at each end, we'd still need to use prediction to compensate for lag when conducting conversations

MORE TO EXPLORE

National Tele-immersion Initiative Web site: www.advanced.org/teleimmersion.html

Tele-immersion at Brown University: www.cs.brown.edu/~lsh/telei.html

Tele-immersion at the University of North Carolina at Chapel Hill: www.cs.unc.edu/Research/stc/teleimmersion/

Tele-immersion at the University of Pennsylvania: www.cis.upenn.edu/~sequence/teleim1.html

Tele-immersion site at Internet2: www.internet2.edu/html/tele-immersion.html

Information about an autostereoscopic display: www.mrl.nyu.edu/projects/autostereo

across the country. This is one reason the current set of test sites are all located on the East Coast.

One promising avenue of exploration in the next few years will be routing tele-immersion processing through remote supercomputer centers in real time to gain access to superior computing power. In this case, a supercomputer will have to be fast enough to compensate for the extra delay caused by the travel time to and from its location.

Bandwidth is a crucial concern. Our demand for bandwidth varies with the scene and application; a more complex scene requires more bandwidth. We can assume that much of the scene, particularly the background walls and such, is unchanging and does not need to be resent with each frame. Conveying a single person at a desk, without the surrounding room, at a slow frame rate of about two frames per second has proved to require around 20 megabits per second but with up to 80-megabit-per-second peaks. With time, however, that number will fall as better compression techniques become established. Each site must receive the streams from all the others, so in a three-way conversation the bandwidth requirement must be multiplied accordingly. The "last mile" of network connection that runs into computer science departments currently tends to be an OC3 line, which can carry 155 megabits per second—just about right for sustaining a three-way conversation at a slow frame rate. But an OC3 line is approximately 100 times more capacious than what is usually considered a broadband connection now, and it is correspondingly more expensive.

I am hopeful that in the coming years we will see a version of tele-immersion that does not require users to wear special glasses or any other devices. Ken Perlin of New York University has developed a prototype of an autostereoscopic display that might make this possible.

Roughly speaking, tele-immersion is about 100 times too expensive to compete with other communications technologies right now and needs more polishing besides. My best guess is that it will be good enough and cheap enough

for limited introduction in approximately five years and for widespread use in around 10 years.

Prospects

WHEN TELE-IMMERSION becomes commonplace, it will probably enable a wide variety of important applications. Teams of engineers might collaborate at great distances on computerized designs for new machines that can be tinkered with as though they were real models on a shared workbench. Archaeologists from around the world might experience being present during a crucial dig. Rarefied experts in building inspection or engine repair might be able to visit locations without losing time to air travel.

In fact, tele-immersion might come to be seen as real competition for air travel—unlike videoconferencing. Although few would claim that tele-immersion will be absolutely as good as "being there" in the near term, it might be good enough for business meetings, professional consultations, training sessions, trade show exhibits and the like. Business travel might be replaced to a significant degree by tele-immersion in 10 years. This is not only because tele-immersion will become

better and cheaper but because air travel will face limits to growth because of safety, land use and environmental concerns.

Tele-immersion might have surprising effects on human relationships and roles. For instance, those who worry about how artists, musicians and authors will make a living as copyrights become harder and harder to enforce (as a result of widespread file copying on the Internet) have often suggested that paid personal appearances are a solution, because personal interaction has more value in the moment than could be reproduced afterward from a file or recording. Tele-immersion could make aesthetic interactions practical and cheap enough to provide a different basis for commerce in the arts. It is worth remembering that before the 20th century, all the arts were interactive. Musicians interacted directly with audience members, as did actors on a stage and poets in a garden. Tele-immersive forms of all these arts that emphasize immediacy, intimacy and personal responsiveness might appear in answer to the crisis in copyright enforcement.

Undoubtedly tele-immersion will pose new challenges as well. Some early users have expressed a concern that tele-immersion exposes too much, that telephones and videoconferencing tools make it easier for participants to control their exposure—to put the phone down or move offscreen. I am hopeful that with experience we will discover both user-interface designs (such as the virtual mirror depicted in the illustration on pages 72 and 73) and conventions of behavior that address such potential problems.

I am often asked if it is frightening to work on new technologies that are likely to have a profound impact on society without being able to know what that impact will be. My answer is that because tele-immersion is fundamentally a tool to help people connect better, the question is really about how optimistic one should be about human nature. I believe that communications technologies increase the opportunities for empathy and thus for moral behavior. Consequently, I am optimistic that whatever role tele-immersion ultimately takes on, it will mostly be for the good. ■

Tele-immersion *Team Members*

■ **UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL:** *Henry Fuchs, Herman Towles, Greg Welch, Wei-Chao Chen, Ruigang Yang, Sang-Uok Kum, Andrew Nashel, Srihari Sukumaran*
www.cs.unc.edu/Research/stc/teleimmersion/

■ **UNIVERSITY OF PENNSYLVANIA**
Ruzena Bajcsy, Kostas Daniilidis, Jane Mulligan, Ibrahim Volkan Isler
www.cis.upenn.edu/~sequence/teleim2.html

■ **BROWN UNIVERSITY**
Andries van Dam, Loring Holden, Robert C. Zeleznik
www.cs.brown.edu/~lsh/telei.html

■ **ADVANCED NETWORKS AND SERVICES**
Jaron Lanier, Amela Sadagic
www.advanced.org/teleimmersion.html

SCIENTIFIC AMERICAN

415 Madison Avenue
New York, NY 10017-1111
www.sciam.com

